

STIC-Biot ch/Ch mLib

61492

From: Pak, Michael  
Sent: Friday, March 01, 2002 8:59 AM  
To: STIC-Biotech/ChemLib  
Subject: 09/276,935 sequence search

Sequence search - 2 month amendment  
App. #: 09/276,935  
Result format: Paper.  
Title: an orphan nuclear receptor

Please search:

Search commercial and interference database.

SEQ ID NO: 13-14

Edward Han  
Technical Info. Specialist  
STIC/Biotech  
MI 6802 Tel: 305-9203

Thanks,

Mike Pak

Michael Pak  
Art Unit 1646  
Mailbox: CM1, Rm. 10C00  
Office: CM1, Rm. 10E13  
703-305-7038

Michael Pak  
USPTO  
Art Unit 1646  
CM1; Rm. 10E13  
703-305-7038

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 3/1/02  
Date Completed: 3/2/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST(where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 84  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



GenoCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2002, 01:50:25 : Search time 22.59 seconds

(without alignments)  
412.411 Million cell updates/sec

Title: US-09-276-935b-14  
Percent score: 2187  
Sequence: 1 LEVREKESMNADPHEEDT.....AMTEFAIPLMGELRGIGS 414

Scoring table:  
Gapop 10.0, Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Lipfold times 45 summaries

Database: Issued patents AA

1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCCTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description                         |
|------------|-------|-------------|-----------|----|-------------------------------------|
| 1          | 759   | 44.3        | 427       | 4  | US-08-764-870-11 Sequence 11, Appl  |
| 2          | 731   | 43.4        | 448       | 1  | US-08-459-488-10 Sequence 10, Appl  |
| 3          | 731   | 43.4        | 448       | 1  | US-08-458-686-10 Sequence 10, Appl  |
| 4          | 731   | 43.4        | 448       | 1  | US-07-843-350C-10 Sequence 10, Appl |
| 5          | 590   | 27.0        | 367       | 1  | US-07-737-736B-4 Sequence 4, Appl   |
| 6          | 472   | 21.6        | 461       | 1  | US-08-430-518-2 Sequence 2, Appl    |
| 7          | 472   | 21.6        | 461       | 1  | US-08-430-284-2 Sequence 2, Appl    |
| 8          | 472   | 21.6        | 461       | 2  | US-08-646-248-2 Sequence 2, Appl    |
| 9          | 472   | 21.6        | 461       | 5  | PCT-US95-13934-2 Sequence 2, Appl   |
| 10         | 472   | 21.6        | 461       | 5  | PCT-US95-13934-2 Sequence 2, Appl   |
| 11         | 470   | 21.5        | 460       | 1  | US-08-342-411A-2 Sequence 2, Appl   |
| 12         | 464   | 21.2        | 446       | 2  | US-08-472-652-3 Sequence 3, Appl    |
| 13         | 464   | 21.2        | 446       | 5  | PCT-US95-16411-4 Sequence 2, Appl   |
| 14         | 461   | 21.1        | 446       | 4  | US-08-776-844-2 Sequence 2, Appl    |
| 15         | 457   | 20.9        | 443       | 1  | US-08-342-411A-4 Sequence 4, Appl   |
| 16         | 456.5 | 20.9        | 440       | 1  | US-08-333-358-8 Sequence 8, Appl    |
| 17         | 456.5 | 20.9        | 440       | 1  | US-08-464-694-8 Sequence 8, Appl    |
| 18         | 456.5 | 20.9        | 440       | 1  | US-08-694-501-8 Sequence 8, Appl    |
| 19         | 456.5 | 20.9        | 447       | 1  | US-08-473-935-1 Sequence 1, Appl    |
| 20         | 424.5 | 19.4        | 472       | 1  | US-08-496-631-2 Sequence 2, Appl    |
| 21         | 423   | 19.3        | 461       | 4  | US-08-764-870-3 Sequence 3, Appl    |
| 22         | 423   | 19.3        | 461       | 4  | US-08-980-115-3 Sequence 3, Appl    |
| 23         | 410   | 18.7        | 455       | 6  | 5223606-4 Patient No. 5223606       |
| 24         | 409   | 18.7        | 463       | 3  | US-08-472-184-2 Sequence 2, Appl    |
| 25         | 409   | 18.7        | 469       | 4  | US-09-469-721-2 Sequence 2, Appl    |
| 26         | 409   | 18.7        | 469       | 5  | PCT-US95-17023-2 Sequence 2, Appl   |
| 27         | 406.5 | 18.6        | 484       | 2  | US-08-372-652-1 Sequence 1, Appl    |

|    |       |      |     |   |                                    |
|----|-------|------|-----|---|------------------------------------|
| 28 | 406.5 | 18.6 | 484 | 5 | PCT-US95-16411-1 Sequence 1, Appl  |
| 29 | 405.5 | 18.5 | 746 | 3 | US-09-144-759-18 Sequence 18, Appl |
| 30 | 405.5 | 18.5 | 746 | 3 | US-09-144-759-20 Sequence 20, Appl |
| 31 | 399.5 | 18.3 | 451 | 2 | US-08-472-652-2 Sequence 2, Appl   |
| 32 | 399.5 | 18.3 | 451 | 5 | PCT-US95-16411-2 Sequence 2, Appl  |
| 33 | 394   | 18.0 | 410 | 4 | US-08-764-870-2 Sequence 2, Appl   |
| 34 | 394   | 18.0 | 410 | 4 | US-08-980-115-2 Sequence 2, Appl   |
| 35 | 394   | 18.0 | 410 | 6 | 5438126-2 Patient No. 5438126      |
| 36 | 390   | 17.8 | 410 | 4 | US-08-764-870-1 Sequence 1, Appl   |
| 37 | 390   | 17.8 | 410 | 4 | US-08-890-115-1 Sequence 1, Appl   |
| 38 | 388   | 17.7 | 606 | 4 | US-08-891-298-3 Sequence 3, Appl   |
| 39 | 387   | 17.7 | 448 | 6 | 5223606-2 Patient No. 5223606      |
| 40 | 379.5 | 17.4 | 433 | 5 | US-08-466-120-2 Sequence 2, Appl   |
| 41 | 379.5 | 17.4 | 433 | 5 | PCT-US94-07260-2 Sequence 2, Appl  |
| 42 | 372   | 17.0 | 462 | 2 | US-08-592-384-2 Sequence 2, Appl   |
| 43 | 372   | 17.0 | 462 | 2 | US-08-095-728B-4 Sequence 4, Appl  |
| 44 | 372   | 17.0 | 462 | 5 | PCT-US92-04200A-4 Sequence 4, Appl |
| 45 | 372   | 17.0 | 797 | 2 | US-08-095-728B-2 Sequence 2, Appl  |

## ALIGNMENTS

RESULT 1  
US-08-764-870-11  
Sequence 11, Applicant App US-08-764870  
Patent No. 6236546  
GENERAL INFORMATION:  
APPLICANT: Scantec, Thomas S  
APPLICANT: Baxter, John D  
APPLICANT: Pfeiffer, Robert J  
APPLICANT: Wagner, Richard L  
APPLICANT: Kushner, Peter J  
APPLICANT: Applebitt, James W  
APPLICANT: West, Brian  
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand  
NUMBER OF INVENTION: 16  
CORRESPONDENCE ADDRESS:  
ADDRESS: Cooley Goddard  
STREET: Five Palo Alto Square, 4000 El Camino Blvd  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US-08-764-870  
FILING DATE: 13-DEC-1996  
CLASSIFICATION: 530  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,540  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,543  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,606  
FILING DATE: 14-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Nakamura, Jackie N  
REGISTRATION NUMBER: 45,966  
REFERENCE/WORK NUMBER: 00-AL-24/0105  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)843-5000  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 427 amino acids  
TYPE: amino acid



NAME: Paul T. Clark  
 REGISTRATION NUMBER: 40,162  
 REFERENCE/WORKET NUMBER: 00786/126001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 TELE: 200154  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 448  
 TYPE: amino acid  
 STRANDEDNESS: N/A  
 TOPOLOGY: linear  
 US-08-458-686-10

Query Match 33.4% Score 731: DB 1: Length 348:  
 Best Local Similarity 40.5%: Pred. No. 1,10-64:  
 Matches 157: Conservative 59: Mismatches 100: Indels 72: Gaps 6:

QY 41 CRYGDKATGHEHNMVTCCKGKFFRRKRNARLCQFRRGACETITKTRPGQACRLR 100  
 DB 11 CVAQGDALGVHFNALTCGCKGKFFRRVSKSIGTQPF-AGSTVEKTKQRHCHVACRLQ 69  
 QY 101 KCLDSGKKKMLMSDAVERRALLKRRKSKRTGVLPGVGLIEGQRMIRIMDAQMK 160  
 DB 70 KCLDSGKKKMLMSDAVERRALLKRRKSKRTGVLPGVGLIEGQRMIRIMDAQMK 123  
 QY 161 TFDTSHEKRNELPGLVLSGCELPESLQAPSRERAAKMSQVRKDLVSLKVSQLRGEGD 220  
 DB 124 HMTDFQVDFQVDFPMALE-IHQ-PLPTLA----- 154  
 QY 221 SVNNYKPAASGCKELFSLPRMAKMSYMKGGIISEAKVISTRDCLIEQVSLKGA 280  
 DB 154 -----VLPVTHFADINTFMVLQVTKTKDLVPSRLIEHQVSLKGA 198  
 QY 281 FEGLQRENFVNAETGIVKGLSYGLIEDIAG-GFQQLLEPMIKRYMKIKGLDEE 339  
 DB 199 VECHIVLNTFQIUNFICPLRYTIEDAKVGFQVPELLEFNHGTIRKQDLEPE 258  
 QY 340 YVLMQALISLSPRGVGLQHRVVDQLEQFATLKSTYIECNRPDPAIRFLTKIMAMT 399  
 DB 259 YVLLAAMALSPRGVGLQHRVVDQLEQFATLKSTYIECNRPDPAIRFLTKIMAMT 318  
 QY 400 PAT-----PIQDEL 408  
 DB 319 LRSINEAVGYQIHTIGLSAMPLQEL 346

RESULT 4  
 US-07-843-3500-10  
 Sequence 10, Application US/07443350C  
 Patent No. 5756448  
 GENERAL INFORMATION:  
 APPLICANT: David D. Moore et al.  
 TITLE OF INVENTION: GAB RECEPTORS AND RELATED  
 NUMBER OF INVENTIONS: MOLECULES AND METHODS  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson  
 REF: 225 Franklin Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: U.S.A.  
 ZIP: 02110 2804  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTERS: IBM OS/2 Model 502 or 55SX  
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
 SOFTWARE: WordPerfect (Version 5.0)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/843-350C  
 FILING DATE: February 26, 1992

CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Paul T. Clark  
 REGISTRATION NUMBER: 40,162  
 REFERENCE/WORKET NUMBER: 00786/126001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 TELE: 200154  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 448  
 TYPE: amino acid  
 STRANDEDNESS: N/A  
 TOPOLOGY: linear  
 US-07-843-3500-10

Query Match 33.4% Score 731: DB 1: Length 348:  
 Best Local Similarity 40.5%: Pred. No. 1,10-64:  
 Matches 157: Conservative 59: Mismatches 100: Indels 72: Gaps 6:

QY 41 CRYGDKATGHEHNMVTCCKGKFFRRKRNARLCQFRRGACETITKTRPGQACRLR 100  
 DB 11 CVAQGDALGVHFNALTCGCKGKFFRRVSKSIGTQPF-AGSTVEKTKQRHCHVACRLQ 69  
 QY 101 KCLDSGKKKMLMSDAVERRALLKRRKSKRTGVLPGVGLIEGQRMIRIMDAQMK 160  
 DB 70 KCLDSGKKKMLMSDAVERRALLKRRKSKRTGVLPGVGLIEGQRMIRIMDAQMK 123  
 QY 161 TFDTSHEKRNELPGLVLSGCELPESLQAPSRERAAKMSQVRKDLVSLKVSQLRGEGD 220  
 DB 124 HMTDFQVDFQVDFPMALE-IHQ-PLPTLA----- 154  
 QY 221 SVNNYKPAASGCKELFSLPRMAKMSYMKGGIISEAKVISTRDCLIEQVSLKGA 280  
 DB 154 -----VLPVTHFADINTFMVLQVTKTKDLVPSRLIEHQVSLKGA 198  
 QY 281 FEGLQRENFVNAETGIVKGLSYGLIEDIAG-GFQQLLEPMIKRYMKIKGLDEE 339  
 DB 199 VECHIVLNTFQIUNFICPLRYTIEDAKVGFQVPELLEFNHGTIRKQDLEPE 258  
 QY 340 YVLMQALISLSPRGVGLQHRVVDQLEQFATLKSTYIECNRPDPAIRFLTKIMAMT 399  
 DB 259 YVLLAAMALSPRGVGLQHRVVDQLEQFATLKSTYIECNRPDPAIRFLTKIMAMT 318  
 QY 400 PAT-----PIQDEL 408  
 DB 319 LRSINEAVGYQIHTIGLSAMPLQEL 346

RESULT 5  
 US-07-737-7368-4  
 Sequence 4, Application US/077377368  
 Patent No. 5260199  
 GENERAL INFORMATION:  
 APPLICANT: Ineluga, Hector F.  
 APPLICANT: Ross, Troy K.  
 APPLICANT: Prahl, Jean M.  
 TITLE OF INVENTION: Method of producing  
 NUMBER OF INVENTIONS: 1,25-Dihydroxyvitamin D<sub>3</sub> Receptor Protein  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Carl R. Schwartz, Esq., c/o Quattris & Brady  
 STREET: 411 East Wisconsin Avenue  
 CITY: Milwaukee  
 STATE: Wisconsin  
 COUNTRY: U.S.A.  
 ZIP: 53202  
 COMPUTER READABLE FORM:



```

QY 359 HAVVQLODFQFATIKSYTECNROPAHRE--LEFKIMAMT-----EFA----- 401
DB 389 FORVEALQDPYVEALISTYTRKRFQDLKRFPMKIKVLSRLISSVHSKQVATRLQOKK 448
QY 402 -TRI-MOELPGI 411
DB 449 LPLLSLIMDV 459

RESULT 7
US-08-830-283-2
Sequence 2, Application US/08830283
Patent No. 5679518
GENERAL INFORMATION:
APPLICANT: Friedman, Elton
APPLICANT: Holmway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Arriei
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
OPERATING SYSTEM: PC-DOS/MS-DOS
FIRMWARE: Patent In Release #1.0, Version #1.25
CLASSIFICATION:
APPLICATION NUMBER: US/08830283
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deidan, Catherine A.
REGISTRATION NUMBER: 46,502
REFERENCE/DRAWING NUMBER: 19327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-830-283-2

Query Match 21.6%; Score 472; DB 1; Length 461;
Best Local Similarity 28.3%; Pred. No. 1e-38;
Matches 122; Conservative 80; Mismatches 119; Indels 110; Gaps 14;
QY 37 GPOICVCGIKATYHVNMTGEGKGFBRAMRNALKCPYR-KCATETTKTRQO 95
DB 83 GHLELWVQIDKRSPPHYVALSTESKGFERRSVVGARRYACGGCTQGMAMFRKQ 142
QY 96 ACHLRKTISSMKKEMISDEAVERKALIKRKSSEPTGTO---PLDVQO----- 142
DB 143 QCLRLKKEASMRQCVLSERQIKKK--IKQOQJESQUSQSPVQPGQSSSSASGPIA 200
QY 143 -----LTHQNMIRKLMQGMKTFPTTTSHERNFRGLRVLSGGE 183
DB 201 SPANSEAGSGSGEGCVQITAAQELMIQOIVAAQIDCKNRKSPS----- 244

```

```

QY 184 LPELSQASHEEPAKWSVCRKIDGLSTKVSQTEKEDKSWNNKPPADSGKETSLLPDM 244
DB 245 -----DQFKVTHWP-----LQAD-----PQSRVAKQKRA---HE 271
QY 244 AUMSTYMEKGIISPAKVISYEPQRPTEHQSILKGAFFHLDENINVENATIGIMWEE-- 401
DB 272 TELATISVQEVIVDAKQVPGQJGHRQVALKASTTEMLLFTAKKNHET---LQTL 428
QY 302 --GRISYCLEP-TANGPOQJLLEPMKFKHYMKKQJLHEEYVLMQALISLSPQKSVLQ 458
DB 329 FIKDFYVSKNDIFBRAGLOVEEINLPFSRPMRIQIDQAFYALIAINIFSAQKINQVE 388
QY 459 HAVVQLODFQFATIKSYTECNROPAHRE--LEFKIMAMT-----EFA----- 401
DB 389 FORVEALQDPYVEALISTYTRKRFQDLKRFPMKIKVLSRLISSVHSKQVATRLQOKK 448
QY 402 -TRI-MOELPGI 411
DB 449 LPLLSLIMDV 459

RESULT 8
US-08-646-248-2
Sequence 2, Application US/08646248
Patent No. 5949422
GENERAL INFORMATION:
APPLICANT: Friedman, Elton
APPLICANT: Holmway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Arriei
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
OPERATING SYSTEM: IBM PC compatible
FIRMWARE: Patent In Release #1.0, Version #1.25
CLASSIFICATION:
APPLICATION NUMBER: US/08646248
FILING DATE: 14-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/440,284
FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Deidan, Catherine A.
REGISTRATION NUMBER: 46,502
REFERENCE/DRAWING NUMBER: 19327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-646-248-2

```







```

0Y 465 LOEFALILKSYTECHRRPOPAIR--ELTKIMMILF-----EFA-----TPLW 406
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
0b 480 LQGVYVALLSYTRIKRPODULFTFRMLMKVSLRTLSVHSRYVAFALRQKRLPLLS 439

0Y 407 EIRPI 411
11 11
1b 440 EIMV 444

RESULT 14
US-08-776-844-2
Sequence 4: Application US/08776844
Patent No. 627976
GENERAL INFORMATION:
APPLICANT: ENMARK, EVA
APPLICANT: GUSTAFSSON, JAN
TITLE OF INVENTION: OR-1 ON ORPHAN RECEPTOR BELONGING
TITLE OF INVENTION: TO THE NUCLEAR RECEPTOR FAMILY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: BARNET & WILCOIT
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZITE: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPIER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,844
FILING DATE: 24-JUN-1997
CLASSIFICATION: 546
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCI/EP95/03247
FILING DATE: 16-AUG-1995
APPLICATION NUMBER: DK 941536.2
FILING DATE: 16-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/CKET NUMBER: 00487, 04029
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 446 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPLOGY: linear
MOLECULE TYPE: protein
OS-08-776-844 2

Query Match 21.18; Score 461; DB 4; Length 446;
Best Local Similarity 28.1%; Prod. No. 1,26-37;
Matches 11%; Conservative 78; Mismatches 12%; Indels 102; Gaps 133

0Y 47 CTGACGCTGCTGATGTAATGTTTGGCGGFFRRAMKKNARLKPFR-KGACETTRKTRRQO 95
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
0b 74 GHECGVGGKAGCAHYNAVISCCKGPKRRSVAGGAGYACGSGTQMDAPMRKQ 133
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

0Y 96 ACRLRKLTSQMKRKMIMSLDAVEERKALIKRKSERTGTPL----- 138
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
0b 144 LCHLRKRLKAGMRQGVVSRHQIRKKRIKUGQDQDQPPPTEDPANSASARAASGISEAS 193
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

0Y 139 -----CGGGLTFGRKRMIRKIMLADQKTDITFSPKPKRLPGVLSGTLPPSLQA 190
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
1b 194 SGGSGSGGCTGCTGTAAGCTMLQGLVAAGCTGNNKRSFS----- 229
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2002, 01:49:40 ; Search time 37.93 Seconds

(without alignments)  
808,498 Million cell updates/sec

File: US-09-276-935b-14  
Perfect score: 2187  
Sequence: 1 LEVPRKSNWNAEVECDL ..... AALFEATPLMDELFGTGS 414

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522464 seqs, 74071290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0.8  
Maximum Match 1.008  
Listing first 45 summaries

Database :

1: /SIDS2/qcqdala/geneseq/geneseq/AA1980.DAT:\*  
2: /SIDS2/qcqdala/geneseq/geneseq/AA1981.DAT:\*  
3: /SIDS2/qcqdala/geneseq/geneseq/AA1982.DAT:\*  
4: /SIDS2/qcqdala/geneseq/geneseq/AA1983.DAT:\*  
5: /SIDS2/qcqdala/geneseq/geneseq/AA1984.DAT:\*  
6: /SIDS2/qcqdala/geneseq/geneseq/AA1985.DAT:\*  
7: /SIDS2/qcqdala/geneseq/geneseq/AA1986.DAT:\*  
8: /SIDS2/qcqdala/geneseq/geneseq/AA1987.DAT:\*  
9: /SIDS2/qcqdala/geneseq/geneseq/AA1988.DAT:\*  
10: /SIDS2/qcqdala/geneseq/geneseq/AA1989.DAT:\*  
11: /SIDS2/qcqdala/geneseq/geneseq/AA1990.DAT:\*  
12: /SIDS2/qcqdala/geneseq/geneseq/AA1991.DAT:\*  
13: /SIDS2/qcqdala/geneseq/geneseq/AA1992.DAT:\*  
14: /SIDS2/qcqdala/geneseq/geneseq/AA1993.DAT:\*  
15: /SIDS2/qcqdala/geneseq/geneseq/AA1994.DAT:\*  
16: /SIDS2/qcqdala/geneseq/geneseq/AA1995.DAT:\*  
17: /SIDS2/qcqdala/geneseq/geneseq/AA1996.DAT:\*  
18: /SIDS2/qcqdala/geneseq/geneseq/AA1997.DAT:\*  
19: /SIDS2/qcqdala/geneseq/geneseq/AA1998.DAT:\*  
20: /SIDS2/qcqdala/geneseq/geneseq/AA1999.DAT:\*  
21: /SIDS2/qcqdala/geneseq/geneseq/AA2000.DAT:\*  
22: /SIDS2/qcqdala/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                 |
|------------|-------|-------------|--------|-------|-----------------------------|
| 1          | 2187  | 100.0       | 414    | 20    | AAV42691 Human pregnane X r |
| 2          | 2167  | 99.1        | 434    | 20    | AAV15941 A human intranucle |
| 3          | 2167  | 99.1        | 457    | 20    | AAV16035 A human intranucle |
| 4          | 2167  | 99.1        | 457    | 20    | AAV15932 A human intranucle |
| 5          | 2167  | 99.1        | 466    | 20    | AAV25410 Human nnk7 partial |
| 6          | 2167  | 99.1        | 473    | 20    | AAV25411 Human nnk7-1 prote |
| 7          | 2167  | 99.1        | 473    | 20    | AAV15936 A human intranucle |
| 8          | 2167  | 99.1        | 473    | 20    | AAV15934 A human intranucle |
| 9          | 2167  | 99.1        | 473    | 20    | AAV09516 Human vitamin D re |
| 10         | 2165  | 99.0        | 437    | 20    | AAV09515 Human vitamin D re |
| 11         | 2049  | 93.7        | 434    | 20    | AAV21799 Human steroid and  |

|    |       |      |     |    |          |                      |
|----|-------|------|-----|----|----------|----------------------|
| 12 | 2049  | 93.7 | 434 | 22 | AAH84417 | Antino acid sequence |
| 13 | 1663  | 66.9 | 316 | 20 | AAV42694 | Hist-prophane X re   |
| 14 | 910.5 | 41.6 | 486 | 17 | AAH98521 | Xenopus cytochrome   |
| 15 | 762   | 34.8 | 423 | 20 | AAW44623 | Kat vitamin D rece   |
| 16 | 759   | 34.7 | 423 | 19 | AAW47509 | Kat vitamin D rece   |
| 17 | 750   | 34.3 | 427 | 20 | AAH81556 | Human vitamin D re   |
| 18 | 750   | 34.3 | 427 | 19 | AAV09064 | Human vitamin D re   |
| 19 | 750   | 34.3 | 427 | 22 | AAH84254 | Antino acid sequence |
| 20 | 750   | 34.3 | 450 | 20 | AAV09036 | Human vitamin D re   |
| 21 | 750   | 34.3 | 477 | 20 | AAV09035 | Human vitamin D re   |
| 22 | 737.5 | 33.7 | 376 | 22 | AAH84255 | Antino acid sequence |
| 23 | 737.5 | 33.4 | 348 | 18 | AAW42556 | Coast (liver) act    |
| 24 | 731   | 33.4 | 348 | 20 | AAW43902 | Human CAR receptor   |
| 25 | 731   | 33.4 | 348 | 22 | AAV63170 | Antino acid sequence |
| 26 | 726   | 33.2 | 348 | 14 | AAH41346 | Human CAR receptor   |
| 27 | 721   | 33.0 | 359 | 22 | AAH63174 | Antino acid sequence |
| 28 | 688.5 | 31.5 | 357 | 20 | AAV17872 | Mouse nuclear rece   |
| 29 | 680   | 31.1 | 358 | 20 | AAW43903 | Mouse CAR receptor   |
| 30 | 680   | 31.1 | 358 | 20 | AAW43903 | Antino acid sequence |
| 31 | 656   | 30.0 | 356 | 19 | AAW47261 | Kat vitamin D rece   |
| 32 | 656   | 30.0 | 356 | 20 | AAW44622 | Kat vitamin D rece   |
| 33 | 593   | 27.1 | 367 | 14 | AAH43656 | Antino acid sequence |
| 34 | 566   | 25.9 | 285 | 22 | AAH63172 | Antino acid sequence |
| 35 | 472   | 21.6 | 460 | 16 | AAH74748 | Human ubiquitons in  |
| 36 | 472   | 21.6 | 461 | 15 | AAH52960 | Human steroid rece   |
| 37 | 472   | 21.6 | 461 | 17 | AAH57982 | Human steroid rece   |
| 38 | 472   | 21.6 | 461 | 17 | AAH57982 | NCK receptor potent  |
| 39 | 470   | 21.5 | 460 | 16 | AAW45034 | Human ubiquitons in  |
| 40 | 466.5 | 21.3 | 445 | 21 | AAV12374 | Mouse CYP1B1. Mu     |
| 41 | 465   | 21.3 | 461 | 17 | AAH46624 | Human foetal Janu    |
| 42 | 464   | 21.2 | 446 | 17 | AAH46624 | Ketoid X recepto     |
| 43 | 461   | 21.1 | 445 | 17 | AAH46624 | Antino acid sequence |
| 44 | 457   | 20.9 | 444 | 16 | AAH74749 | Kat ubiquitons in    |
| 45 | 457   | 20.9 | 444 | 18 | AAW45035 | Kat ubiquitons in    |

## ALIGNMENTS

|          |   |
|----------|---|
| RESULT 1 |   |
| AAV42691 | Standard: Protein: 414 AA.  |
| ID       | AAV42691  |
| XX       |   |
| AC       | AAV42691  |
| XX       |   |
| DT       | 17-JAN-2000 (first entry)   |
| XX       |   |
| DE       | Human pregnane X receptor (hpxr).                                       |
| XX       |   |
| KW       | Human; nuclear receptor; pregnane X receptor; EPR; CYP; CYP4A4;         |
| KW       | cytochrome P-450 mono-oxygenase; drug interaction; hpxr.                |
| XX       |   |
| OS       | Homo sapiens.   |
| XX       |   |
| PN       | W09948915-A1.   |
| XX       |   |
| PD       | 30-SEP-1999.  |
| XX       |   |
| XX       | 26-MAR-1999; 99WO-0506737.  |
| XX       |   |
| PR       | 27-MAR-1998; 98US-0079593.  |
| XX       |   |
| PA       | (GLAXO ) GILXO GROUP LTD.   |
| XX       |   |
| PI       | Kilmer SA, Willison   |
| XX       |   |
| DP       | WP1: 1999-601202/51.  |
| DK       | N-PSOH: AAV07997.   |
| XX       |   |
| PT       | New human pregnane X receptor, used to identify specific modulators and |
| XX       | agents that induce expression of cytochrome P-450 mono-oxygenase        |
| PS       | Claim 4: Fig 1A-D: 64pp; English.                                       |

*Bad date*

*5500 people*

XX The invention provides an isolated human nucleot receptor (designated  
 XX productive X receptor, RXR) that binds to a retinoid, e.g. 9-*trans*-octadecahydro-  
 XX (2E)-phenanthrene-10-one (Vitamin A), and is used to identify: (a) its specificity for  
 XX and compounds that induce RXR expression; (b) to identify drug  
 XX and compounds that induce RXR expression; (c) to identify RXR  
 XX targets. The nucleot receptor (RXR) is involved in many biological functions of  
 XX drugs. The nucleot receptors are potential targets for association with  
 XX diseases and conditions with RXR and for treatment of such conditions.  
 XX Antisense raised against RXR can be used for determination and  
 XX purification of RXR. The present sequence represents the RXR.

XX Sequence: 414 AA:

Query Match: 100.0% Score: 2187; pos 20; Length: 414;  
 Best Local Similarity: 100.0% Pos: 5; to 209;  
 Mismatches: 444; Conserved: 0; Mismatches: 0; Indels: 0; Gaps: 0;

UY 1 LEVYKRSNMAHVEHTEESVQKPSVNAHVEVQVTEVQVQKATVHNVNMTES 60  
 LB 1 LEVYKRSNMAHVEHTEESVQKPSVNAHVEVQVTEVQVQKATVHNVNMTES 60  
 UY 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 LB 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 UY 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 LB 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 UY 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 LB 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 UY 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 LB 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 UY 181 CTEPESVQASREFAAKWSQVKQKQKSTKQVQKQKQKQKQKQKQKQKQKQKQK 240  
 LB 181 CTEPESVQASREFAAKWSQVKQKQKSTKQVQKQKQKQKQKQKQKQKQKQKQK 240  
 UY 241 PHMAQNSTYMEKGISVAKVISTYKQVQKQKQKQKQKQKQKQKQKQKQKQK 400  
 LB 241 PHMAQNSTYMEKGISVAKVISTYKQVQKQKQKQKQKQKQKQKQKQKQKQK 400  
 UY 401 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 LB 401 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 UY 461 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 LB 461 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 UY 461 VVQVQK 460  
 LB 461 VVQVQK 460

RESULT 2  
 AAY16045 standard: Protein: 434 AA.

XX AAY16045.  
 XX 04 AUG 1999 (first entry)

XX A human intranuclear receptor protein.

XX Human intranuclear receptor protein; drug development; diagnostic;  
 KW treatment.

XX Homo sapiens.

XX 011112/672 A

XX 18 MAY 1999

XX 07 AUG 1999 98DP 0224172

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

XX N PDB: AAY16045.

XX New intranuclear receptor protein used for drug development and  
 XX diagnosis and treatment of disease.

XX Claim 1: Page 15, 16; 48pp; Japanese.

XX The present sequence represents a human intranuclear receptor protein  
 XX the nucleot acid sequence was isolated from a human adult cDNA  
 XX library using a specific AN24 derived probe. The protein can  
 XX be used for the development of drugs and diagnostics and treatment  
 XX of various diseases.

XX Sequence: 414 AA:

Query Match: 99.1% Score: 2187; pos 20; Length: 414;  
 Best Local Similarity: 99.1% Pos: 5; to 209;  
 Mismatches: 444; Conserved: 0; Mismatches: 0; Indels: 20; Gaps: 1;

UY 1 LEVYKRSNMAHVEHTEESVQKPSVNAHVEVQVTEVQVQKATVHNVNMTES 60  
 LB 1 LEVYKRSNMAHVEHTEESVQKPSVNAHVEVQVTEVQVQKATVHNVNMTES 60  
 UY 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 LB 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 UY 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 LB 61 CKETPRARKRNARKEFERKCAETFEKTRGOCVAVRQKTESQKKEKIMSDAVHE 120  
 UY 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 LB 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 UY 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 LB 121 KRALIKRKKSEKTEGIVQVGLTEEGQKMLKEIMQAKKEETPSHEKNEPGLVSS 180  
 UY 181 CTEPESVQASREFAAKWSQVKQKQKSTKQVQKQKQKQKQKQKQKQKQKQKQK 240  
 LB 181 CTEPESVQASREFAAKWSQVKQKQKSTKQVQKQKQKQKQKQKQKQKQKQKQK 240  
 UY 241 PHMAQNSTYMEKGISVAKVISTYKQVQKQKQKQKQKQKQKQKQKQKQKQK 400  
 LB 241 PHMAQNSTYMEKGISVAKVISTYKQVQKQKQKQKQKQKQKQKQKQKQKQK 400  
 UY 401 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 LB 401 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 UY 461 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 LB 461 CPELSYCEKIVAKVQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 460  
 UY 461 VVQVQK 460  
 LB 461 VVQVQK 460

RESULT 3  
 AAY16045 standard: Protein: 457 AA.

XX AAY16045.

XX 04 AUG 1999 (first entry)

XX A human intranuclear receptor protein.

XX Human intranuclear receptor protein; drug development; diagnostic;  
 KW treatment.

XX Homo sapiens.

XX 011112/672 A

XX 18 MAY 1999

XX 07 AUG 1999 98DP 0224172

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

XX 11 AUG 1999 97DP 0240400

```

XX 11-AUG-1997: 97JP-0230335.
XX (NISH) JAPAN TOHACCO INC.
XX WPI: 1999-350330/40.
XX N-PSDB: AAX59975.
XX New intranuclear receptor protein - useful for drug development and
XX diagnosis and treatment of disease.
XX Disclosure: Page 35-37: 38pp: Japanese.
XX T: present sequence represents a human intranuclear receptor protein.
XX The nucleic acid sequence was isolated from a human adult cDNA
XX library using a swollish ANO23 derived probe. The protein can
XX be used for the development of drugs and diagnosis and treatment
XX of various diseases.
XX Sequence 457 AA:

Query Match: 99.18% Score 2167; DB 20; Length 457;
Best Local Similarity: 95.48% Pred. No. 60 207; Indels 20; Gaps 1;
Matches 414; Conservative 0; Mismatches 0;

YY 1 LEVNRKESNMHADPHVCHDFESVQKPSVNADEVGSDVQICRGVQKATYHNNVMTCES 60
DB 24 LEVPRKESNMHADPHVCHDFESVQKPSVNADEVGSDVQICRGVQKATYHNNVMTCES 84
YY 61 CGGFEFRANKRNARKLCRPERKACETLRTKRGCGVCRLEKLESCMKKEMMSDEAVHE 120
DB 64 CGGFEFRANKRNARKLCRPERKACETLRTKRGCGVCRLEKLESCMKKEMMSDEAVHE 143
YY 121 RRALIKRKRSEKRTGTOPLGVGGLTHEORNMHEIMDAOMKTPPTFSHFKNRFGVLS 180
DB 144 RRALIKRKRSEKRTGTOPLGVGGLTHEORNMHEIMDAOMKTPPTFSHFKNRFGVLS 203
YY 181 GCELPESLQASREKAAKMSQVRKDCSLKVSQICRGRDSVNNYKRPANSSKELFSSL 240
DB 204 GCELPESLQASREKAAKMSQVRKDCSLKVSQICRGRDSVNNYKRPANSSKELFSSL 263
YY 241 PHMAKSTYMEKGLISPAKVISYERDLPLEQISLKGAFAELQLEPNTVFNAETGWE 300
DB 264 PHMAKSTYMEKGLISPAKVISYERDLPLEQISLKGAFAELQLEPNTVFNAETGWE 323
YY 301 CGRLSVLELTAAGCPQOGLLEPMKFEHMLKKLOHREHYVMQALSLSPDRGVYQHR 360
DB 324 CGRLSVLELTAAGCPQOGLLEPMKFEHMLKKLOHREHYVMQALSLSPDRGVYQHR 383
YY 361 VVDLOQDFATILKSYTECNQVPAHRTFLKIMAMLE-----F 400
DB 384 VVDQIQGFATILKSYTECNQVPAHRTFLKIMAMLELTSIHAGHTGTLTIGDHP 443
YY 401 ATPMOLEFGTGS 414
DB 444 ATPMOLEFGTGS 457

```

```

XX JP1112782-A.
XX 18-MAY-1999.
XX WPI: 1999-350330/30.
XX N-PSDB: AAX59967.
XX 07-AUG-1998: 98JP-0224172.
XX 11-AUG-1997: 97JP-0230335.
XX (NISH) JAPAN TOHACCO INC.
XX WPI: 1999-350330/30.
XX N-PSDB: AAX59967.
XX New intranuclear receptor protein - useful for drug development and
XX diagnosis and treatment of disease.
XX Claim 2: Page 16-17: 38pp: Japanese.
XX The present sequence represents a human intranuclear receptor protein.
XX The nucleic acid sequence was isolated from a human adult cDNA
XX library using a swollish ANO23 derived probe. The protein can
XX be used for the development of drugs and diagnosis and treatment
XX of various diseases.
XX Sequence 457 AA:

Query Match: 99.18% Score 2167; DB 20; Length 457;
Best Local Similarity: 95.48% Pred. No. 60 207; Indels 20; Gaps 1;
Matches 414; Conservative 0; Mismatches 0;

YY 1 LEVNRKESNMHADPHVCHDFESVQKPSVNADEVGSDVQICRGVQKATYHNNVMTCES 60
DB 24 LEVPRKESNMHADPHVCHDFESVQKPSVNADEVGSDVQICRGVQKATYHNNVMTCES 84
YY 61 CGGFEFRANKRNARKLCRPERKACETLRTKRGCGVCRLEKLESCMKKEMMSDEAVHE 120
DB 84 CGGFEFRANKRNARKLCRPERKACETLRTKRGCGVCRLEKLESCMKKEMMSDEAVHE 143
YY 121 RRALIKRKRSEKRTGTOPLGVGGLTHEORNMHEIMDAOMKTPPTFSHFKNRFGVLS 180
DB 144 RRALIKRKRSEKRTGTOPLGVGGLTHEORNMHEIMDAOMKTPPTFSHFKNRFGVLS 203
YY 181 GCELPESLQASREKAAKMSQVRKDCSLKVSQICRGRDSVNNYKRPANSSKELFSSL 240
DB 204 GCELPESLQASREKAAKMSQVRKDCSLKVSQICRGRDSVNNYKRPANSSKELFSSL 263
YY 241 PHMAKSTYMEKGLISPAKVISYERDLPLEQISLKGAFAELQLEPNTVFNAETGWE 300
DB 264 PHMAKSTYMEKGLISPAKVISYERDLPLEQISLKGAFAELQLEPNTVFNAETGWE 323
YY 301 CGRLSVLELTAAGCPQOGLLEPMKFEHMLKKLOHREHYVMQALSLSPDRGVYQHR 360
DB 324 CGRLSVLELTAAGCPQOGLLEPMKFEHMLKKLOHREHYVMQALSLSPDRGVYQHR 383
YY 361 VVDLOQDFATILKSYTECNQVPAHRTFLKIMAMLE-----F 400
DB 384 VVDQIQGFATILKSYTECNQVPAHRTFLKIMAMLELTSIHAGHTGTLTIGDHP 443
YY 401 ATPMOLEFGTGS 414
DB 444 ATPMOLEFGTGS 457

```



QY 301 CGRLSYGLHAGFOOGLLEPMLKPHMLKKLQJHEEYVLMQALSLSPDRQVYQHR 360  
 DB 340 carlsyglactaaqqlqlllepmklhymkkklqllheeyvlmqalslspdrpvylqhr 399  
 QY 461 VVQJLOFQFALIKSYTEENRQPHRELFLIKIMAMTE-----F 400  
 DB 400 vvdqjqfqtatllksyteenrppahrltllkImamtelrstaahltqrltrigqthpr 459  
 QY 401 ATPMOELFETGS 414  
 DB 460 atpmlqeltdtgs 473

## RESULT 7

AAV15936  
 ID AAV15936 standard: protein: 473 AA.

AAV15936:

04-AUG-1999 (first entry)

XX A human intranuclear receptor protein.

DE Human: intranuclear receptor protein: drug development; diagnosis;  
 treatment.

XX ~~Human: sapiens.~~

XX JP11127872-A.

XX 18-MAY-1999.

XX 07-AUG-1998; 98JP-0224172.

XX 11-AUG-1997; 97JP-0230335.

XX (NISHI) JAPAN TOHACCO INC.

XX WPI: 1999-350330/40.

XX N-PSDB: AAX99974.

XX New intranuclear receptor protein - useful for drug development and  
 diagnosis and treatment of disease

XX Disclousure: Page 32-35; 38pp; Japanese.

XX The present sequence represents a human intranuclear receptor protein.

XX The nucleic acid sequence was isolated from a human adult cDNA  
 library using a swedish ANO23 derived probe. The protein can  
 be used for the development of drugs and diagnosis and treatment  
 of various diseases.

XX Sequence 473 AA:

Query Match 99.1%; Score 2167; DB 20; Length 473;  
 Best Local Similarity 95.4%; Pred. No. 6,30-207;  
 Matches 414; conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 LEVPRKESMHAIPVHCEDESVKSPKSYNAHVEGGOICWCGKALCYHFNVMTCG 50  
 DB 40 levprksswhadlvncedevspkpsvnahevqqltrvvdkaayhfnvmrcq 99  
 QY 61 CKGFRRAMKRNALRCYPRKACGLTKTRKQVACRLKRLCSGKRKEMISDAVEK 120  
 DB 100 ckgtrrramkrnarlrctprkracgltktrqcacltrkclsgmkkemidsavee 159  
 QY 121 RRALIKRKSRKSTQFLVQGLTEPQRMATREIMDAQMKTFTTFSHFRRPLPGVLS 180  
 DB 160 rralikrkssrtqflvqqlteeqrmmitrindaqmktlfttshfkrltprylss 219  
 QY 181 GCLPESLQAPSRHAAKMSQVKKDLASLQSLQRLGDSVWYKPPAISGKELIFSL 240

DB 220 gcelpselqapsrhaakmsqvrkdlcslkvsllqrlgdsavwykppadsqkeltsl 279  
 QY 241 RHMAQDSMTWPKETLSFAKVISYFROLPLEQJSLKGAFFELVQMFNIVNAETGWE 300  
 DB 280 rhmaqdsmtwktlsfakvisyftdpleqjstlkgaafelvtmfnvnaetgwe 339  
 QY 301 CGRLSYGLHAGFOOGLLEPMLKPHMLKKLQJHEEYVLMQALSLSPDRQVYQHR 360  
 DB 340 carlsyglactaaqqlqlllepmklhymkkklqllheeyvlmqalslspdrpvylqhr 399  
 QY 461 VVQJLOFQFALIKSYTEENRQPHRELFLIKIMAMTE-----F 400  
 DB 400 vvdqjqfqtatllksyteenrppahrltllkImamtelrstaahltqrltrigqthpr 459  
 QY 401 ATPMOELFETGS 414  
 DB 460 atpmlqeltdtgs 473

## RESULT 8

AAV15943  
 ID AAV15943 standard: protein: 473 AA.

AAV15943:

04-AUG-1999 (first entry)

XX A human intranuclear receptor protein.

DE Human: intranuclear receptor protein: drug development; diagnosis;  
 treatment.

XX ~~Human: sapiens.~~

XX JP11127872-A.

XX 18-MAY-1999.

XX 07-AUG-1998; 98JP-0224172.

XX 11-AUG-1997; 97JP-0230335.

XX (NISHI) JAPAN TOHACCO INC.

XX WPI: 1999-350331/40.

XX N-PSDB: AAX99998

XX New intranuclear receptor protein - useful for drug development and  
 diagnosis and treatment of disease

XX Claim 2: Page 17-19; 38pp; Japanese.

XX The present sequence represents a human intranuclear receptor protein.

XX The nucleic acid sequence was isolated from a human adult cDNA  
 library using a swedish ANO23 derived probe. The protein can  
 be used for the development of drugs and diagnosis and treatment  
 of various diseases.

XX Sequence 473 AA:

Query Match 99.1%; Score 2167; DB 20; Length 473;  
 Best Local Similarity 95.4%; Pred. No. 6,30-207;  
 Matches 414; conservative 0; Mismatches 0; Indels 20; Gaps 1;

QY 1 LEVPRKESMHAIPVHCEDESVKSPKSYNAHVEGGOICWCGKALCYHFNVMTCG 60  
 DB 40 levprksswhadlvncedevspkpsvnahevqqltrvvdkaayhfnvmrcq 99  
 QY 61 CKGFRRAMKRNALRCYPRKACGLTKTRKQVACRLKRLCSGKRKEMISDAVEK 120  
 DB 100 ckgtrrramkrnarlrctprkracgltktrqcacltrkclsgmkkemidsavee 159



PR 41-MAR-1998; 98SEP-0001145  
PR 14-OCT-1997; 97SEP-0003745

PR 14-0011-1997; 97SH-0003745

PA (FIMA) PHARMACIA &amp; TJ.P.JOHIN AB.

PI HURKISTAM A, Dahlberg M

WP1: 1999-302508/25.

IXR N-PSIB; AAX<sup>56242</sup>.

PI New vitamin D receptor related (VDR) polypeptides, useful for treating obesity, diabetes, anorexia and rheumatoid arthritis

PS Claim 19; Page 19-20; 35pp; English.

The present sequence is a human vitamin D receptor related (VDR) polypeptide. Human VDR polypeptides and substances which affect VDR signal transduction, can be used for treating metabolic, proliferative or inflammatory conditions. They can be used in the manufacture of a medicament for treating the following conditions: obesity, diabetes, hypocalcaemia, lipoprotein defects, hyperlipidaemia, hypercholesterolaemia or hypertriglyceridaemia and osteoporosis, rheumatoid arthritis, benign and malignant tumours, hyperproliferative skin disorders or hyperthyroidism. Nucleic acid vectors encoding for expression of a VDR polypeptide can be used for treating metabolic, proliferative or inflammatory conditions by introducing them into a mammal. The introduced nucleic acid is then capable of transforming a cell *in vivo* and then polypeptide is expressed. A substance affecting VDR signal transduction, such as an agonist or antagonist can be used for the manufacture of a medicament for treating metabolic, proliferative or inflammatory condition.

N.B. The specification specifically claims the VDR nucleic acid and polypeptide sequences given in figures 1, 4, 7 and 8, but no figures are given in the specification.

Sequence 447 AA:

Query Match (94.0%) Score 2165; DH 20; Length 4372

Best local similarity: 95.2% pred. No. 8,9e-207;  
Matches: 413; Conserved: 0; Indels: 20; Gaps: 1

```

QY 1 LEKPESSNNAAVEHCEITSEVSGKTSVNADEFGADPQICVCGIKATGYPFNMYTCGG 60
Db 1 merriykesmhadihcedtesvpgkpsvnaadevgqpietvceqdkatqyhtnmiceg 60
QY 61 CKGFFFRAMKRNAPLCPEFKAGCEITRKTRPOVQACRLKLCESGKKKMTSTFAVEE 120
Db 61 CKGATIRAKMRAITCPDFKAGCEITRKTRPGQACRIKLCESGKKKMTSTFAVEE 120
QY 121 RRATIKRKKSPKRGUPLATGACITTEQDPMHRLFMADKMTPTTFPSHKHNRDQVLS 180
Db 121 RRAIKRKKSPKRGUPLATGACITTEQDPMHRLFMADKMTPTTFPSHKHNRDQVLS 180
QY 181 GCELPKSLAPSKERFAAMSVQRKDIATLKVSLQTRQEDSVWVKKYPADSGKPESTL 240
Db 181 GCELPKSLAPSKERFAAMSVQRKDIATLKVSLQTRQEDSVWVKKYPADSGKPESTL 240
QY 241 PHAAQKSTYKFKGITSFKAVISYFEDRLIEVQSLKLGAAFEICQLRENTVNAETGWE 300
Db 241 PHAAQKSTYKFKGITSFKAVISYFEDRLIEVQSLKLGAAFEICQLRENTVNAETGWE 300
QY 361 CGRLSTCEITFMAGCVQDILIEMLKRFHYMLKQIDHEPEYVLMGALSTSPORQGVQHR 420
Db 361 CGRLSTCEITFMAGCVQDILIEMLKRFHYMLKQIDHEPEYVLMGALSTSPORQGVQHR 420
QY 481 VVDQIDQFAITLKSYTFCNRIDQAHRLPLKIMAMLT-----F 400
Db 481 VVDQIDQFAITLKSYTFCNRIDQAHRLPLKIMAMLT-----F 400
QY 401 ATPLMQELHETGGS 414
Db 421 ATPLMQELHETGGS 414

```

## RESULTS

MAY 21 7 59 AM '66

AAV21799 standard: Protein; 434 AA

AAV217992

DT 14-SHIP-1999 (first entry)

Human steroid and xenobiotic receptor (SXR),

KW Nuclear receptor; SXK, steroid and xenobiotic receptor; RXK, human;

phytoestrogen

polycystic ovary

(15) Homo sapiens

[illegible]

|    |         |
|----|---------|
| FT | unknown |
| FT | unknown |

XXXXXX

1000

XX  
XXXX

XX  
XX

XX  
XX  
P / Conference 15

XX 1001-1000-41  
XX 1001-1000-41

IR N-15103; AAX8  
XX

| PT | New steroid |
|----|-------------|
| DT | 2004-2011   |

PT their toxic

PS Claim 4; Fig  
XY

The invention relates to a novel nuclear receptor polypeptide, designated SXR (steroid and xenobiotic receptor). SXR (i) forms a heterodimer with retinoid X receptor (RXR), (ii) binds to a direct or inverted repeat response element motif based on the half-site AATTA, (iii) activates transcription through response elements present in steroid-inducible p450 genes, in response to a wide variety of natural and synthetic steroid hormones and (iv) is prominently expressed in liver and intestine. SXR regulates expression of catalytic enzymes, in response to many different steroids, and thus affects metabolism. SXR is a broad specificity, low affinity receptor for reducing excessive levels of steroids in the circulation. (Anti)agonists of SXR are used to regulate metabolism of steroids particularly phytoestrogens or calcimimetic blockers, to reduce steroid toxicity in subjects being treated with steroids, e.g., in cases of tuberculosis (treated with rifampin and related compounds), breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis (treated with Vitamin K), or to slow metabolism of therapeutic steroids. Also, modulating endogenous SXR is used to treat disease, particularly an agonist is used where endogenous steroid levels are excessive (e.g., Cushing's syndrome, virilism and hirsutism in women; polycystic ovarian disease). It has a 17- or 21-hydroxylase deficiency, 3 beta-hydroxysteroid dehydrogenase deficiency, or breast, colorectal or prostate cancer). SXR while antagonists are used where endogenous steroid levels are too low. Cells that express SXR are used to identify compounds likely to be involved in immunohistochemical reactions. Antibodies specific for SXR are used in immunohistochemical testing for studying distribution/expression density of SXR, also for diagnosis and therapeutically as antagonist the present sequence represents SXR polypeptide.



XX His5-pregnane X receptor (PXR) sequence.  
 CE  
 XX Human nuclear receptor: pregnane X receptor; PXR; CYP3A4;  
 KW cytochrome P-450 mono-oxygenase; drug interaction; hpxr.  
 XX  
 XX Synthetic  
 XX Homo Sapiens  
 XX W09948915-A1  
 XX 30-SEP-1999  
 XX 26-MAR-1999: 95W0-US006737  
 XX 27-MAR-1998: 98US-0074594.  
 XX (GLAXO) GLAXO GROUP LTD.  
 XX KILBEAR SA, WILLIS TOWERS  
 XX WPI: 1999-601202/51.  
 XX New human pregnane X receptor, used to identify specific modulators and  
 XX agents that induce expression of cytochrome P-450 mono-oxygenase  
 XX Example 6: Page 37: 6pp; English.  
 XX The invention provides an isolated human nuclear receptor (designated  
 XX pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase  
 XX (CYP) promoter. The hpxr is used to identify its specific modulators,  
 XX and compounds that induce CYP3A4 expression (i.e., to identify drug  
 XX interactions, since CYP3A4 is involved in many biotransformations of  
 XX drugs). The modulators are potentially useful for: associating particular  
 XX diseases and conditions with PXR and for treating such conditions.  
 XX Antibodies raised against hpxr can be used for determination and  
 XX purification of hpxr. The present sequence represents a histidine-6  
 XX tagged partial PXR (His6-PXR) sequence.  
 XX  
 XX Sequence 416 AA:

Query Match 46.98; Score 1463; DB 20; Length 316;  
 Best Local Similarity 93.48; Pred. No. 4e-137;  
 Matches 285; Conservative 0; Mismatches 0; Indels 20; Gaps 1;

UY 140 SERGIGPVUGLTFEFGMMHELMAGMKETFTPSHFKNFRIPVLSGCELPSTLQ 189  
 DB 12 SRRGQPIGVQGLTFEFGMMHELMAGMKETFTPSHFKNFRIPVLSGCELPSTLQ 71  
 UY 190 APSRFAAKSVQWKDGLSKVSLQIRHDSVWYKPPASGCKETPSLPHMAQMSY 249  
 DB 72 APSRFAAKSVQWKDGLSKVSLQIRHDSVWYKPPASGCKETPSLPHMAQMSY 131  
 UY 250 MFKLISFAVIVYFRIPIEQSILKGAETQLQFNIVFAETITWAGLSTGLE 309  
 DB 112 MFKLISFAVIVYFRIPIEQSILKGAETQLQFNIVFAETITWAGLSTGLE 191  
 UY 310 DIALNFOULLFEMKPHVMKRIQLHEEYVLMQALSLFSDRCVGLHFRVYQLQEP 369  
 DB 142 DIAHQVQLLEPMKIKHMKIKLQHEEYVLMQALSLFSDRCVGLHFRVYQLQEP 251  
 UY 370 AITLAKSTFERNKVPARIRFLKTMALTE-----FATVLMQLP 409  
 DB 252 AITLAKSTFERNKVPARIRFLKTMALTE-----FATVLMQLP 311  
 UY 410 GILGS 414  
 DB 312 GILGS 316

RESULT 14  
 AAR98521

ID AAR98521 standard: Protein: 386 AA.  
 AC AAR98521;  
 XX  
 XX 14-NOV-1996 (first entry)  
 DE Xenopus orphan receptor 6.  
 XX  
 XX Xenopus orphan receptor 6; XOR 6; steroid receptor; vitamin D,  
 KW hydroxycortisol receptor; metoprolol; aminobenzoate.  
 XX  
 XX Xenopus laevis.  
 XX  
 XX Key location/qualifiers  
 FT Domain 37-102  
 FT Domain 148-386  
 FT Domain 148-386  
 XX W09622390-A1.  
 XX 25-JUL-1996.  
 XX 16-JAN-1996: 96W0-US00058.  
 XX 17-JAN-1995: 95US-0374445.  
 XX (SALK) SALK INST BIOLOGICAL STUDIES.  
 XX Blumberg B, Evans RM, Imamoto K;  
 XX WPI: 1996-354546/35.  
 XX N-PSDB: AAT36499.  
 XX DNA encoding receptor polypeptide responsive to hydroxy, metoprolol  
 XX and amino benzoate(s) - useful to regulate gene transcription  
 XX Claim 4: Page 27-28: 4pp; English.  
 XX  
 XX Xenopus orphan receptor 6 (XOR-6) (AAR98521) is a new member of the  
 XX steroid receptor superfamily, characterized as being responsive to  
 XX the presence of hydroxy, metoprolol or amino benzoate(s) and as  
 XX regulating the transcription of associated genes(s). It shows 78%  
 XX identity in the DNA binding domain, and 42% identity in the ligand  
 XX binding domain, to the human vitamin D receptor. Recombinant XOR-6  
 XX can be expressed in animal cells; a cDNA clone (AAT36499) coding for  
 XX XOR-6 has been isolated. The recombinant XOR-6 may be used to  
 XX regulate gene transcription or to raise antibodies of diagnostic  
 XX or therapeutic appln.  
 XX  
 XX Sequence 386 AA:

Query Match 41.68; Score 910.5; DB 17; Length 386;  
 Best Local Similarity 46.28; Pred. No. 5.1e-942;  
 Matches 192; Conservative 55; Mismatches 100; Indels 69; Gaps 10;

UY 18 EDPESVNGKSVNADVEGVPQIRVQGHKATGVHNVMTQVSKQFRKAMRNARLKC 77  
 DB 14 EECQDASNSVQVQEDDDDDPKFTRACDADLQYHNMCCVCEKATIRAVYRLLSC 74  
 UY 78 PFKKAGELTETKTRKQVQCHQIRKCLSSGKKEMKSLAEVFRALIRK-KSEFTG 146  
 DB 74 PFKKAGELTETKTRKQVQCHQIRKCLSSGKKEMKSLAEVFRALIRK-KSEFTG 142  
 UY 137 PLVVGGLTFEFGMMHELMAGMKETFTPSHFKNFRIPVLSGCELPSTLQ 196  
 DB 133 PLVVGGLTFEFGMMHELMAGMKETFTPSHFKNFRIPVLSGCELPSTLQ 171  
 UY 197 AKMSQVGRKRLSLKVSLOIRKGGDSVWYKPPASGCKETPSLPHMAQMSYFKE 256  
 DB 172 AKMSQVGRKRLSLKVSLOIRKGGDSVWYKPPASGCKETPSLPHMAQMSYFKE 207







Genoto version 4.5  
Copyright (c) 1994 - 2000 Computen Ltd.

OM protein protein search, using SW model

Run on: March 2, 2002, 01:53:20 : Search time 33.2 Seconds

(without alignments)  
949,887 Million cell updates/sec

Title: US-09-276-935b-14

Perfect score: 2187

Sequence: 1 LEVPEKESWNIADPVHCPD.....AMLEFATPIIMQELPITGS 414

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR,68:\*

1: PIR,68:\*

2: PIR,68:\*

3: PIR,68:\*

4: PIR,68:\*

5: PIR,68:\*

6: PIR,68:\*

7: PIR,68:\*

8: PIR,68:\*

9: PIR,68:\*

10: PIR,68:\*

11: PIR,68:\*

12: PIR,68:\*

13: PIR,68:\*

14: PIR,68:\*

15: PIR,68:\*

16: PIR,68:\*

17: PIR,68:\*

18: PIR,68:\*

19: PIR,68:\*

20: PIR,68:\*

21: PIR,68:\*

22: PIR,68:\*

23: PIR,68:\*

24: PIR,68:\*

25: PIR,68:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match length | DB ID        | Description        |
|------------|-------|--------------------|--------------|--------------------|
| 1          | 922.5 | 42.2               | 386 2 J04497 | thyroid hormone re |
| 2          | 922.5 | 42.2               | 388 2 J07510 | benzoate X recepto |
| 3          | 922.5 | 42.2               | 420 2 J07229 | vitamin D receptor |
| 4          | 922.5 | 42.2               | 446 2 J04571 | vitamin D receptor |
| 5          | 922.5 | 42.2               | 423 2 A11761 | 1,25-dihydroxyvita |
| 6          | 922.5 | 42.2               | 427 2 J04019 | vitamin D receptor |
| 7          | 922.5 | 42.2               | 427 2 A28200 | vitamin D receptor |
| 8          | 922.5 | 42.2               | 427 2 S21174 | vitamin D receptor |
| 9          | 922.5 | 42.2               | 425 2 J07230 | vitamin D receptor |
| 10         | 922.5 | 42.2               | 448 2 A56197 | nuclear hormone re |
| 11         | 922.5 | 42.2               | 446 2 J04014 | steroid hormone re |
| 12         | 922.5 | 42.2               | 446 2 J04021 | retinoid X recepto |
| 13         | 922.5 | 42.2               | 445 2 A56043 | steroid hormone re |
| 14         | 922.5 | 42.2               | 446 2 A56043 | orphan nuclear rec |
| 15         | 922.5 | 42.2               | 447 2 J08975 | nuclear orphan rec |
| 16         | 922.5 | 42.2               | 369 2 J08975 | thyroid hormone re |
| 17         | 922.5 | 42.2               | 369 2 S58211 | beta-thyroid hormo |
| 18         | 922.5 | 42.2               | 373 2 C56067 | thyroid hormone re |
| 19         | 922.5 | 42.2               | 473 2 C56067 | thyroid hormone re |
| 20         | 922.5 | 42.2               | 456 2 J04019 | thyroid hormone re |
| 21         | 922.5 | 42.2               | 373 2 J04019 | thyroid hormone re |
| 22         | 922.5 | 42.2               | 461 2 A40377 | thyroid hormone re |
| 23         | 922.5 | 42.2               | 461 2 A11820 | thyroid hormone re |
| 24         | 922.5 | 42.2               | 475 2 A57035 | thyroid hormone re |
| 25         | 922.5 | 42.2               | 478 2 A41055 | ecdysone receptor  |
| 26         | 922.5 | 42.2               | 456 2 A56590 | ecdysone receptor  |
| 27         | 922.5 | 42.2               | 459 2 A56918 | retinoid X recept  |
| 28         | 922.5 | 42.2               | 484 2 J04018 | thyroid hormone re |
| 29         | 922.5 | 42.2               | 416 2 J01096 | thyroid hormone re |

## ALIGNMENTS

| Query Match           | Score | 42.2%   | Score    | 922.5%     | DB 2% | Length | 48% |
|-----------------------|-------|---|----------|------------|-------|--------|-----|
| Best Local Similarity | 46.6% | Prod. No.   | 6.4e-65% |            |       |        |     |
| Matches               | 194   | Conservative  | 54       | Mismatches | 99    | Indels | 69  |
| Gaps                  | 10    |   |          |            |       |        |     |
| Query                 | 18    | EDTESVGRKSVNADDEFGVQICVGDGKATGSHFVNTGKCFEHRAMKRLNC        | 77       |            |       |        |     |
| DB                    | 14    | EEEDASNSGICEDBEDDQKICACGDGKATGSHFVNTGKCFEHRAMKRLNC        | 73       |            |       |        |     |
| Query                 | 78    | FEKGCATLTKLIRQCACRLKCI ESQMKKEMIMSDAVERRALIKK-KSERICGV    | 146      |            |       |        |     |
| DB                    | 74    | FEKGCATLTKLIRQCACRLKCI ESQMKKEMIMSDAVERRALIKK-KSERICGV    | 142      |            |       |        |     |
| Query                 | 137   | PLVCGGLFEDGKMLKELMDQMKCFDTESSPKKMLPVLSGGLPSPSLQAS-REFA    | 195      |            |       |        |     |
| DB                    | 133   | PLVCGGLFEDGKMLKELMDQMKCFDTESSPKKMLPVLSGGLPSPSLQAS-REFA    | 171      |            |       |        |     |
| Query                 | 197   | AKMSQVKKDLSKVSLSQIKGDSVWNTKPPVDSGCKELSLIPMAIMSTYMKGLIS    | 256      |            |       |        |     |
| DB                    | 172   | AKMSQVKKDLSKVSLSQIKGDSVWNTKPPVDSGCKELSLIPMAIMSTYMKGLIS    | 207      |            |       |        |     |
| Query                 | 257   | FAKVSIFRRLPLFEDQISLKGAAFEICQLRPVFNANFGTMTCCGSLNGIEDT-AGSP | 315      |            |       |        |     |
| DB                    | 208   | FAKVSIFRRLPLFEDQISLKGAAFEICQLRPVFNANFGTMTCCGSLNGIEDT-AGSP | 267      |            |       |        |     |
| Query                 | 316   | QQLLEPMKLFHYMLKKQIAEEFYVLMQALSLSPHPVVLQRRVVLQLOFALILKS    | 375      |            |       |        |     |
| DB                    | 268   | QQLLEPMKLFHYMLKKQIAEEFYVLMQALSLSPHPVVLQRRVVLQLOFALILKS    | 327      |            |       |        |     |
| Query                 | 376   | YIEKNC-PQAIHFLKLMAMLEF-----ATPLMGLFG 410                  |          |            |       |        |     |
| DB                    | 328   | YIEKNC-PQAIHFLKLMAMLEF-----ATPLMGLFG 410                  |          |            |       |        |     |







Query 191 PSREHAKKNSQVHKDLSKVS-----LQLEKSDNSVWNYKPPAUSGKEIFESLLPIKADW 246  
 Db 179 LLSASSDSFSHSESV-DIKVNFNNLMWYDFGSS-----SPHSREHGGSSSMPLPIADL 213  
 247 STYWFKEIISPAVIVSYFRLDPLEDLSLKGAFAELCOLNFNVFAFGTMCOCG--L 304  
 Db 214 VYVSLQVQVIGFAMIGFELTAEQDIALKSSALEVIMLSNSFNLEFMSKSGADPD 293  
 305 SVYLEED-TAGFQOALLFEMKLFHMLKIQLEHEEVYVMAISLSPDRQVYLRKVD 163  
 Db 294 KYVLSIVIKAGRTLELEPLVKKVQVLTAKKINQHEEHVLMALICLLSPDRQVQDARLE 153  
 364 QLOQFMTIKSYIEGNRPQPAHRELEKIMAMLEF----- 400  
 Db 354 ALQURKSTGATQILQHL-HPQGRITVAKMQLADLRSLNEHNSKQYRSISREHRSKO 411  
 401 APLMEELFG 410  
 Db 412 LTPLELVSG 421

## RESULT 10

AS6197  
 nuclear hormone receptor MB67 - human  
 C1Species: Homo sapiens (man)  
 C1Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 20-Sep-1999  
 C1Accession: A56197  
 C1Bases: M: Gullick, T: Chou, H S: Martinoli, M.G.: Sima, D.: Moore, D.D.  
 Mol. Cell. Biol. 14: 1544-1552, 1994  
 A1Title: A new orphan member of the nuclear hormone receptor superfamily that interacts  
 A1Reference number: A56197; M01D:94158627  
 A1Accession: A56197  
 A1Status: preliminary  
 A1Molecule type: mRNA  
 A1Residues: 1-348 <BAES>  
 A1Cross-references: GH:230425; GH:129263; NID:9458541; PUDN:CAAB3016.1; PTD:q458542  
 C1Superfamily: unassigned erba-related proteins; erba transforming protein homology  
 C1Keywords: zinc finger  
 459-270/Domain: erba transforming protein homology <ERBA>

Query Match 33.4%; Score 711; DB 2; Length 348;  
 First Local Similarity 40.5%; Pred. No. 6, 1e-50;  
 Matches 157; Conservative 59; Mismatches 100; Indels 72; Gaps 6;

Query 41 CRVGCDATGYNHNVMTGCKCKPFRKAKKNAKLRGCFKGCAGETTRKTRQOACRLR 100  
 Db 11 GVGCGDQATGHNALTCGCKGFFRIVSKSTGTPCP-AGSEVSKTORRRCPCARLQ 69  
 101 KTEESKKEMKNSLEAVERKRLAKKKSEKRTGTPYVGLTFEUGMMIRELMQMK 160  
 Db 70 KGLDARKKMILISAFALIRAKAKQNGRAQOTPVG-----LSKQGEIKETLSAHTR 123  
 161 TTDITSEHFNHFGVLSNGTELPESLDAPSRFAAKWSYQRDLVSLKYSQLRDEG 220  
 Db 124 HMGTFEDFVGVRPAALFTHD-PLPTLAP----- 153  
 221 SWNYKPPAUSGKEIFESL PHMAKMSIYFKGIIISPAKVISYFRLDPLEDQSLKGA 280  
 Db 154 -----VPLVTHFAQNTNFMVLQVLEKTKDLVERSLPTEQDLSLKGA 198  
 281 FEIDQLNFNVNMTSTWFGSLSGCEDTAG-GEQOALLEPMKPKHYMKKQJLHEE 319  
 Db 159 VETGHVINTFELQTNELQGLKRTTFDQAKVQVYAFPLLFHFGITKRLQIDQEP 258  
 340 YVLMQALISFSPDRQVYLRKVDYQLEQAFATIKSYIEGNRPQPAHRELEKIMAM 399  
 Db 259 YVLLAAMALFSPDRQVYLRKVDLEQLEQAFATIKSYIEGNRPQPAHRELEKIMAM 318  
 400 FAT -----PLMEEL 408  
 Db 419 LRSLNEAVQVQLQISLSSAMPLIQL 346

## RESULT 11

JC4014  
 steroid hormone-nuclear receptor NER - human  
 C1Species: Homo sapiens (man)  
 C1Date: 13-Jun-1995 #sequence\_revision 14-Jul-1995 #text\_change 24-Sep-1999  
 C1Accession: JC4014  
 R1Shinar, D.M.; Endo, N.; Rutledge, S.J.; Voyle, R.; Kodan, G.A.; Schmidt, A.  
 Gene 147, 273-276, 1994  
 A1Title: NER, a new member of the gene family encoding the human steroid hormone nucl  
 A1Reference number: JC4014; M01D:95011628  
 A1Accession: JC4014  
 A1Molecule type: mRNA  
 A1Residues: 1461 <SHN>  
 A1Cross-references: GH:007142; NID:q641961; PUDN:AAA61784.1; PTD:q641962  
 A1Experimental source: osteosarcoma cells SMO-2/HU  
 C1Genetics:  
 A1Gene: GDB:UNK  
 A1Cross-references: GDB:489270; OMIM:600480  
 A1Map position: 19q13.4-19q13.3  
 C1Superfamily: unassigned erba-related proteins; erba transforming protein homology  
 C1Keywords: steroid hormone receptor  
 F185-381/Domain: erba transforming protein homology <ERBA>  
 F187-154/Domain: DNA binding #status predicted <BIN>

Query Match 21.6%; Score 472; DB 2; Length 461;

First Local Similarity 28.3%; Pred. No. 1, 9e-29;  
 Matches 122; Conservative 80; Mismatches 119; Indels 110; Gaps 14;

Query 37 GPQICRVGCDKATGYNVTCGCKPFRKAKKNAKLRPER-KGAEITRTKRCQV 95  
 Db 83 GHLCVCGIKAGCPHYVNLSCGCKGFFRIVSKSGAKRYACGRDTCOMQAFMRKCV 142  
 96 ACRLKRTKSGKKEMKNSLEAVERKRLAKKKSEKRTGTPYVGLTFEUGMMIRELMQMK 142  
 Db 143 QCHLRKKEKEMKNSLEAVERKRLAKKKSEKRTGTPYVGLTFEUGMMIRELMQMK 200  
 143 -----LTFEUGMMIRELMQMKITPTTSHTKRNRLQVYSSGTE 183  
 Db 201 SPQSPAGSGQSGCGVGLVAQRLMLOHVAADLPNKRSP----- 244  
 184 LPSIADPSHFAKNSQVYKRLQSLKYSQLEKSDNSVWNYKPPAUSGKEIFESLLIM 243  
 Db 245 -----LQPKVTHMP-----LQAD-----PQSRDAQDQFA--HF 271  
 244 ADMSTYFKGIIISPAVIVSYFRLDPLEDLSLKGAFAELCOLNFNVFAFGTMCOCG-- 401  
 Db 272 TELALISVEELVDFKQVPGFQIGREQDIALKASTELMLETARKYHET---ECIT 428  
 302 --GRTSGVLEED-TAGFQOALLFEMKLFHMLKIQLEHEEVYVMAISLSPDRQVY 458  
 Db 329 FIKPFTYSKHFRAQVQVLEFNTLTFEFSKMRHRIQIDNATVALLAINTFSADHNVQ 388  
 359 HRYVDQLQEQAFATIKSYIEGNRPQPAHRELEKIMAMLEF-----EFA----- 401  
 Db 389 PGVFAALQVYVFAALSYTRIKRQDQLRPRMCKIVSLTSSVSHQVFAKIQMK 448  
 402 -TPLMEELG 411  
 Db 449 LPELSTSLMW 459

## RESULT 12

149021  
 retinoid X receptor interacting protein No.15 - mouse  
 C1Species: Mus musculus (house mouse)  
 C1Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Sep-1999  
 C1Accession: 149021  
 R1Seol, W.; Choi, H.S.; Moore, D.D.  
 Mol. Endocrinol. 9, 72-85, 1995  
 A1Title: Isolation of proteins that interact specifically with the retinoid X recepto  
 A1Reference number: A57664; M01D:95280959



[illegible]

**RESULT** 15  
 138475  
 nuclear orphan receptor LXR-alpha - human  
 C:Species: Homo sapiens (man)  
 C:date: 16-Feb-1996 #sequence\_revision 16-Feb-1996 #text\_change 20-Sep-1999  
 C:Accession: 138475  
 R:Willy, P.-J.; Oleson, K.; Ong, E.-S.; Evans, R.M.; Heyman, R.A.; Mangelsdorf, D.J.  
 Genes Dev. 9, 1043-1045, 1995  
 A:Title: LXR, a nuclear receptor that defines a distinct retinoid response pathway  
 A:Reference number: 138975; MUIP:95262897  
 A:Accession: 138975  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1447 - RES-  
 A:Cross-references: EMBL:U2662; NID:9726512; PDB:AAA9856.1; PDB:9726513  
 C:Superfamily: unassigned eTRA-related proteins; eTRA transforming protein homology  
 C:Keywords: zinc finger  
 E:96-367/domain: eTRA transforming protein homology <EKKA>

```

QY 407 GLEETA-GSGUULLLEPMKRFYMKKJOLHFEFVYVMOA1SIFSPRQYVJOHVVDOL 408
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 422 NKRETFKAGLGVFFTPD1FERSKANE1QJNDAPFALTA1S1FSADKRONVQJLVYERL 423
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 366 QEOPALFLKSYTCRNPQFAHFLFLK1KIMALTFFAL-----1LM 405
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 382 QHTYVEMLMNVYS1HHFH--DRMLTPHMLKLV1K1TSSVNSEQVAFK1QJDKK1P1L 409
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 406 QELPGI 411
      | : | :
Db 440 SETMDV 445

```



GenInfo version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 2, 2002, 02:37:30 ; Search time 24.95 seconds

(without alignments)  
633,789 Million cell updates/sec

Title: US-09-276-935b-14

Perfect score: 2187  
Sequence: 1 LEVPEKESNMHADVHCHDT.....AMLEPAPLMQHLPGITGS 414

Scoring table: HUSUM62  
Gapop 10.0 ; Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Testing first 45 summaries

Database: SwissProt\_39\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 2165   | 94.0        | 434    | 1     | FXR_HUMAN   |
| 2          | 1675.5 | 76.6        | 431    | 1     | FXR_MOUSE   |
| 3          | 1656.5 | 75.7        | 431    | 1     | FXR_RAT     |
| 4          | 774    | 35.4        | 422    | 1     | VDR_XENLA   |
| 5          | 768.5  | 35.1        | 448    | 1     | VDR_COTIA   |
| 6          | 768    | 35.1        | 451    | 1     | VDR_CHICK   |
| 7          | 762    | 34.8        | 423    | 1     | VDR_CAT     |
| 8          | 751.5  | 34.5        | 422    | 1     | VDR_MOUSE   |
| 9          | 750    | 34.3        | 424    | 1     | VDR_BOVIN   |
| 10         | 731    | 34.3        | 427    | 1     | VDR_HUMAN   |
| 11         | 685.5  | 31.3        | 348    | 1     | NR13_HUMAN  |
| 12         | 680    | 31.1        | 358    | 1     | NR13_RAT    |
| 13         | 680    | 31.1        | 358    | 1     | NR13_MOUSE  |
| 14         | 472    | 21.6        | 441    | 1     | NRH2_HUMAN  |
| 15         | 469.5  | 21.5        | 445    | 1     | NRH2_MOUSE  |
| 16         | 463    | 21.2        | 446    | 1     | NRH2_MOUSE  |
| 17         | 461.5  | 21.1        | 445    | 1     | NRH2_RAT    |
| 18         | 461    | 21.1        | 446    | 1     | NRH2_RAT    |
| 19         | 456.5  | 20.9        | 447    | 1     | NRH3_HUMAN  |
| 20         | 433    | 19.8        | 386    | 1     | THB_BRARE   |
| 21         | 431.5  | 19.7        | 395    | 1     | THB_PAROL   |
| 22         | 428.5  | 19.6        | 757    | 1     | ETR_LICCU   |
| 23         | 426    | 19.5        | 369    | 1     | THB_CHICK   |
| 24         | 426    | 19.5        | 461    | 1     | THB1_HUMAN  |
| 25         | 425    | 19.4        | 373    | 1     | THB_XENLA   |
| 26         | 424    | 19.4        | 373    | 1     | THB1_XENLA  |
| 27         | 423.5  | 19.4        | 675    | 1     | ECR_AEIAE   |
| 28         | 423    | 19.3        | 416    | 1     | THA_HIPPI   |
| 29         | 423    | 19.3        | 476    | 1     | THB2_HUMAN  |
| 30         | 419    | 19.2        | 414    | 1     | THB1_XENLA  |
| 31         | 417    | 19.1        | 411    | 1     | THB1_SHEEP  |
| 32         | 417    | 19.1        | 461    | 1     | THB1_MOUSE  |
| 33         | 417    | 19.1        | 475    | 1     | THB2_MOUSE  |

|    |       |      |     |   |            |                    |
|----|-------|------|-----|---|------------|--------------------|
| 34 | 413.5 | 18.9 | 878 | 1 | ECR_DROME  | P44622 diophylla   |
| 35 | 413   | 18.9 | 461 | 1 | THB1_RAT   | P18113 ratius nov  |
| 36 | 413   | 18.9 | 514 | 1 | THB2_RAT   | P37626 ratius nov  |
| 37 | 411.5 | 18.8 | 427 | 1 | THB1_BRARE | Q98862 cytochrome  |
| 38 | 411.5 | 18.8 | 536 | 1 | ECR_CHILE  | P49882 cytochrome  |
| 39 | 406   | 18.6 | 416 | 1 | THA_PAROL  | Q91241 parolichthy |
| 40 | 402.5 | 18.4 | 556 | 1 | ECR_MOUSE  | P49883 mardica sex |
| 41 | 401.5 | 18.4 | 408 | 1 | THA_CHICK  | P46625 galus gall  |
| 42 | 397.5 | 18.2 | 402 | 1 | THA_APTPA  | Q42295 aptinodytes |
| 43 | 397.5 | 18.2 | 402 | 1 | THA_PYGAD  | Q42450 pygostolus  |
| 44 | 395.5 | 18.1 | 401 | 1 | THA_CAIMO  | Q96482 caimomus    |
| 45 | 394   | 18.0 | 410 | 1 | THA1_SHEEP | Q28570 ovis aries  |

## ALIGNMENTS

| RESULT | 1  | STANDARD: | PRT: | 434 AA. |
|--------|--|-----------|------|---------|
| ID     | FXR_HUMAN  |           |      |         |
| AC     | Q75469:  |           |      |         |
| DT     | 15-JUL-1999 (Ref. 38, Created)   |           |      |         |
| DT     | 15-JUL-1999 (Ref. 78, Last sequence update)  |           |      |         |
| DT     | 20-AUG-2001 (Ref. 40, Last annotation update)  |           |      |         |
| DE     | ORPHAN NUCLEAR RECEPTOR FXR (PREGNANE X RECEPTOR)  |           |      |         |
| DE     | RECEPTOR PAR1).  |           |      |         |
| GN     | NR112 OR FXR   |           |      |         |
| OS     | Homo sapiens (Human).  |           |      |         |
| OC     | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  |           |      |         |
| OC     | Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  |           |      |         |
| OX     | NCBI_TaxId:9606;   |           |      |         |
| RP     | (1)  |           |      |         |
| RP     | SEQUENCE FROM N.A.   |           |      |         |
| RC     | TISSUE: Liver;   |           |      |         |
| FX     | MEDLINE:98495174; PubMed:9727070;  |           |      |         |
| RA     | Lehmann J.M., McKee D.D., Watson M.A., Willson T.M., Moore J.L.,   |           |      |         |
| RA     | Kliwer S.A.;   |           |      |         |
| RT     | *The human orphan nuclear receptor FXR is activated by compounds that  |           |      |         |
| RT     | regulate CYP24 gene expression and cause drug interactions.*;  |           |      |         |
| RL     | J. Clin. Invest. 102:1016-1023(1998).  |           |      |         |
| RN     | (2)  |           |      |         |
| RP     | SEQUENCE FROM N.A.   |           |      |         |
| RC     | TISSUE: Liver;   |           |      |         |
| FX     | MEDLINE:98445450; PubMed:970465;   |           |      |         |
| RA     | Bertilsson G., Hultich J., Swenson K., Asman M., Jendberg L.,  |           |      |         |
| RA     | Sydvall-Hackman M., Olsson K., Postlund H., Blomquist P.,  |           |      |         |
| RA     | Bertensham A.;   |           |      |         |
| RT     | *Identification of a human nuclear receptor defines a new signaling  |           |      |         |
| RT     | pathway for CYP1A induction.*;   |           |      |         |
| RL     | Proc. Natl. Acad. Sci. U.S.A. 95:12208-12213(1998).  |           |      |         |
| CC     | -1- FUNCTION: ORPHAN RECEPTOR. ITS NATURAL LIGAND IS PROBABLY  |           |      |         |
| CC     | PREGNANE. BINDS TO A RESPONSE ELEMENT IN THE CYP1A4 GENE PROMOTER.   |           |      |         |
| CC     | -1- SUPPLEMENT: FORMS A HETERODIMER WITH FXR.  |           |      |         |
| CC     | -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).   |           |      |         |
| CC     | -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, COLON, AND SMALL   |           |      |         |
| CC     | INTESTINE.   |           |      |         |
| CC     | -1- INDUCTION: ACTIVATED BY NATURALLY OCCURRING STEROIDS SUCH AS   |           |      |         |
| CC     | PREGNENOLONE AND PROGESTERONE.   |           |      |         |
| CC     | -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.   |           |      |         |
| CC     | NR1 SUBFAMILY.   |           |      |         |
| CC     | This SWISS-PROT entry is copyright. It is produced through a collaboration                                       |           |      |         |
| CC     | between the Swiss Institute of Bioinformatics and the EMBL outstation.   |           |      |         |
| CC     | use by non-profit institutions as long as its content is in no way   |           |      |         |
| CC     | modified and this statement is not removed. Usage by and for commercial  |           |      |         |
| CC     | entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a> |           |      |         |
| CC     | or send an email to <a href="mailto:license@isb.slb.ch">license@isb.slb.ch</a> ).                                |           |      |         |
| DR     | EMBL: AF001054; AAC05436.1;  |           |      |         |
| DR     | EMBL: AF084645; AAC04558.1;  |           |      |         |
| DR     | EMBL: AF084645; AAC04558.1;  |           |      |         |
| DR     | EMBL: AF084645; AAC04558.1;  |           |      |         |
| DR     | InterPro: IPR000536; Hormone_rec_fam.  |           |      |         |

| PRODUCT | EXR MODEL#   | STANDARD | FRZ | 431 AA |
|---------|--|----------|-----|--------|
| 10      | EXR MOOSE  |          |     |        |
| AC      | 06/01/97   |          |     |        |
| DT      | 15, 01, 1999 (6-1, 48, 12/01/00)                             |          |     |        |
| 11      | 15, 01, 1999 (6-1, 48, last sequence updated)                |          |     |        |
| 12      | 20 Apr 2001 (6-1, 40, last annual report update)             |          |     |        |
| 13      | COPIED INTO LEAR REFERENCE EXR (PRESERVE X REFERENCE)        |          |     |        |
| 14      | NR12, 01R EXR  |          |     |        |
| 15      | MOB INVENTORIES (MOBILE)                                     |          |     |        |
| 16      | BOOKING/INFORMATION/CHORDING/STANDARD/VOICED/RECORDING/      |          |     |        |
| 17      | MEMORIAL/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/  |          |     |        |
| 18      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 19      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 20      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 21      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 22      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 23      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 24      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 25      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 26      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 27      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 28      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 29      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 30      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 31      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 32      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 33      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 34      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 35      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 36      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 37      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 38      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 39      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 40      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 41      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 42      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 43      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 44      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 45      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 46      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 47      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 48      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 49      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 50      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 51      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 52      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 53      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 54      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 55      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 56      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 57      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 58      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 59      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 60      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 61      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 62      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 63      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 64      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 65      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 66      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 67      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 68      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 69      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 70      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 71      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 72      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 73      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 74      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 75      | RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/RECORDING/ |          |     |        |
| 76      | RECORDING/RECORDING/RECORDING/RE                             |          |     |        |

|     |     |               |                  |        |             |            |      |       |          |
|-----|-----|---------------|------------------|--------|-------------|------------|------|-------|----------|
| QY  | 4   | VPRKSSWALPVEH | EHFESVIRKISNAQDE | EVQVQV | YR          | QVQKATVY   | ENWV | FEQR  | 6        |
| 10b | 1   | MPHESWQVETVQV | TEAUSALREH       | 1      | WVEFHESQVQV | QVWADKANVY | YH   | NYWV  | FEQR     |
| QY  | 63  | QYHREHAKKAKK  | PEPRKQV          | EE     | IKRIRKQV    | QVQV       | TE   | IKRQV | ESKQKKEH |
| 10b | 60  | QYPRQAKKBNKLE | TPPRKQV          | TE     | IKRIRKQV    | QVQV       | TE   | IKRQV | ESKQKKEH |
| QY  | 123 | ALIKRKSSEKQV  | QVQV             | TE     | IKRIRKQV    | QVQV       | TE   | IKRQV | ESKQKKEH |
| 10b | 120 | ALIKRKSSEKQV  | QVQV             | TE     | IKRIRKQV    | QVQV       | TE   | IKRQV | ESKQKKEH |
| QY  | 184 | ELPESQVQV     | SHQVAAKSSQV      | QVQV   | TE          | IKRIRKQV   | QVQV | TE    | IKRQV    |
| 10b | 180 | ELPESQVQV     | SHQVAAKSSQV      | QVQV   | TE          | IKRIRKQV   | QVQV | TE    | IKRQV    |
| QY  | 243 | MAHNSYHKEH    | SLKAVSY          | TE     | IKRIRKQV    | QVQV       | TE   | IKRQV | ESKQKKEH |

DB 240 LADVSTYMKGVINFAKVI/SYFROLPLEHQSILKCAIPEMCTIKRNTWEDIFIGWCG 299

QY 403 RLSVCLIEDIAGGPOQLLEPMLEKHYMKLQJLHEEYVLMQATISFSDRQVQLHRVY 362

DB 400 RLAVCFEDNGGFGKLLDLPMLKREHCKMLKQLDLKEEYVLMQATISFSDRQVQSVV 359

QY 363 DOJGQFATLTKSYTECKNPQAHFPLFKIMAMLE-----FAI 402

DB 360 DOJGQFATLTKAVIFCSKYPAPHFPLFKIMAVLELKSINMQQTLRIQDSHPAFI 419

QY 404 PLMGELRGIT 412

DB 420 PLMGELRGIT 429

RESULT 3

PXR\_RAT 3 STANDARD: PRT: 431 AA.

AC G91A7:

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE 20-AUG-2001 (Rel. 40, Last annotation update)

DE ORPHAN NUCLEAR RECEPTOR PXR (PREGNANE X RECEPTOR).

CN NR12 OR PXR.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN 1

RP SEQUENCE FROM N.A.

KX MEDLINE=99345883; PubMed=10415105;

KA Zhang H., Lechlyse E., Liu L., Lu M., McInerney L., Zhu W., Yan B.;

RT "Rat pregnane X receptor: molecular cloning, tissue distribution, and xenobiotic regulation."

RL Arch. Biochem. Biophys. 398:14-22(1999).

CC -1- FUNCTION: ORPHAN RECEPTOR. ITS NATURAL LIGAND IS PROBABLY PREGNANE. BINDS TO A RESPONSE ELEMENT IN CYPA GENES PROMOTER (BY SIMILARITY).

CC -1- SUBUNIT: FORMS A HETERODIMER WITH RXR (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY. NR1 SUBFAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL consortium - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC

DB EMBL: AF151377; A0474214.1.

DB InterPro: IPR000536; hormone\_rec\_lig.

DB InterPro: IPR001723; steroidhormone\_rptfor.

DB InterPro: IPR001628; zt-c4.

DB Pfam: PF00104; hormone\_rec\_1.

DB Pfam: PF00105; zt-c4\_1.

DB PRINTS: PR00947; STEROIDRECEP.

DB PRINTS: PR00350; VITAMINDR.

DB PRINTS: PR00398; STEROHORMONR.

DB SMART: SM00430; HDL1.1.

DB SMART: SM00339; ZNF\_C4\_1.

DB PROSITE: PS00041; NUCLEARRECEPTOR\_1.

KM Receptor; Transcription regulation; DNA-binding; Nuclear protein;

KW Zinc-finger.

FW DNA\_BIND 38 104 NUCLEAR RECEPTOR-TYPE.

FT ZN\_FING 38 58 C4-TYPE.

FT ZN\_FING 74 99 C4-TYPE.

FT DOMAIN 105 201 HINGE.

FT DOMAIN 202 431 LIGAND-BINDING.

SO SEQUENCE 431 AA. 49660 MW. 4854521F949697 CRC64;

Query Match 76.7% Score 1656.5 DB 1; Length 431;

Best Local Similarity 72.8%; Prod. No. 1.70e-126;

Matches 433; Conservative 41; Mismatches 55; Indels 21; Gaps 2;

QY 3 VHKKSWNHADPVHCHDIDIESYCKPSNADPEHGGVQICRNVADKATGVHVMICDQK 62

DB 1 MPEEFKNNHVGIVGRFADSVLEER-INVDEIDGQICRNVQDRKNDGVHVMKIQKAK 59

QY 63 GEFRRAMKRNALRCPFKKACETIRKIRRCVACRIRKCTESKREKIMSDFAVEER 122

DB 60 GEFRRAMKRNVALRCPFKKICETIRKIRRCVACRIRKCTESKREKIMSDFAVEER 119

QY 123 ALIKRKKRIRICIGPLISVQGLIEGQRMILREIMACKRTPTDIESKPKRPLVLSGG 182

DB 120 ALIKRKKRIRIRPAPPGQGLIEGQALIDPMQADQVITPTFSFKPLKPLAVSGG 179

QY 183 ELPEISQVSRFEARKWSQVRKDIQSLKVSQGLRQHSQVNYKRPASAKNPLPSLPR 242

DB 180 ELPEVQASLLEDPATWSQIMKDSVPMKISVQLRGDLSIMNYPKSKKELIPILPH 239

QY 243 MAMSTYMKKGISFAKVISFQDLPEIDQSLIKKAAPEQLRNTVNMFTWNG 402

DB 240 LADVSTYMKGVINFAKVI/SYFROLPLEHQSILKCAITPMCTIKRNTWEDIFIGWCG 299

QY 403 RLSVCLIEDIAGGPOQLLEPMLEKHYMKLQJLHEEYVLMQATISFSDRQVQLHRVY 462

DB 400 RLAVCFEDNGGFGKLLDLPMLKREHCKMLKQLDLKEEYVLMQATISFSDRQVQSVV 459

QY 363 DOJGQFATLTKSYTECKNPQAHFPLFKIMAMLE-----FAI 402

DB 360 DOJGQFATLTKAVIFCSKYPAPHFPLFKIMAVLELKSINMQQTLRIQDSHPAFI 419

QY 404 PLMGELRGIT 412

DB 420 PLMGELRGIT 429

RESULT 4

VDR\_XENIA 4 STANDARD: PRT: 422 AA.

ID VDR\_XENIA

AC 013124;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).

CN VDR OR NR111.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mosobatrachia; Pipridae; Pipidae;

OC Xenopodidae; Xenopus.

OX NCBI\_TaxID=8355;

RN 111

RP SEQUENCE FROM N.A.

KX TISSUE-KIDNEY: PubMed=9165021;

KX MEDLINE=97407679; PubMed=9165021;

KA LA Y.C., Bernawitz G., Jaeppe H., Demay M.B.;

RT Cloning and characterization of the vitamin D receptor from Xenopus laevis.

RL Endocrinology 138:2347-2353(1997).

CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING THE EXPRESSION OF HORMONE-SENSITIVE GENES.

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- TISSUE SPECIFICITY: DETECTED IN ALL THE TISSUE EXAMINED. HIGHEST LEVEL IN SMALL INTESTINE AND SKIN.

CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT STAGE 14, INCREASES GRADUALLY AND PEAKS AT STAGE 57-61 THEN DECREASES TO THE LEVEL SEEN IN ADULT.

CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY. NR1 SUBFAMILY.

CC -----



```

0b 1MELIADIVSVLSIOVKYFAKMLPGCFROLDIAEDVADIALKSSALFVIMLRNSOSFTWEDMS 306
0y 299 MEGR--LASCLED--FAGSGQGLILEPMLKRFHYMKKLOLHEEYVYMAISLFSFDRIG 355
Db 407 WIGOSNPQKKKQVADYVQASHSMOLLEPLVYKFOVGLKKRIINLEPFLVILMAITLSPDRG 366
0y 456 VLAHVVQVQVQGFATLEKSYLETCNKPQAHKLEPLKIMATLEF----- 400
Db 467 VQDTSLVESIQDMLSDLTQYTRCHREPPGSRLLYAKMIGIADLRSLNEHSKYRCIS 426
0y 401 -----ATFLMQLG 410
Db 427 KQPHSMQPLVLEVFG 444

RESULT 5
VDR_CHICK
10 VDR_CHICK STANDARD: PRF: 45) AA.
AC 062342;
DE 15-DEC-1998 (Rel. 37, Created)
DE 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (L25-DIHYDROXYVITAMIN D3 RECEPTOR).
DE VDR OR NR11.
DS Gallus gallus (Chicken).
DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eupholostomi;
DE Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
DE Gallus.
DE NCBI_TaxID: 9041;
0x NCBI_TaxID: 9041;
11
RP SEQUENCE FROM N.A.
RP STRAIN LEHJOHN; TISSUE: Kidney;
RP MEDLINE-97223369; PubMed-9056239;
RA Lu Z., Hanson K., DeLuca H.F.;
RA "Cloning and origin of the two forms of chicken vitamin D receptor.*";
RL Arch. Biochem. Biophys. 339:99-106(1997).
12
RP SEQUENCE OF 45-114 FROM N.A.
RP MEDLINE 87149040; PubMed-1029866;
RA McConnell D.P., Mangelsdorf D.J., Pike J.W., Haussler M.R.;
RA O'Malley B.W.;
RA "Molecular cloning of complementary DNA encoding the avian receptor
RT for vitamin D*";
RL Science 235:1214-1217(1987).
CC
CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) ARE PRODUCED BY USE OF
CC ALTERNATIVE INITIATION CODONS.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY AND INTESTINE.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NOT SUBMITTABLE.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF011456; AAB62579.1;
0R InterPro: IPR000546; Hormone_rec_119.
0R InterPro: IPR016281; Z1-C4.
0R Pfam: PF00104; hormone_rec_1.
0R Pfam: PF00105; z1-c4.1
0R PRINTS: PR00047; STROIDEINR.
0R PRINTS: PR00450; VITAMINDR.
0R SMART: SM00450; b01.1.
0R SMART: SM00450; znf_C4.1.

```

```

KW PROSITE (PS00041): NUCLEAR RECEPTOR 1.
KM Receptor, Transcription regulation, DNA-binding, Nuclear protein.
KW Zinc-finger.
FT CHAIN 1 451 VITAMIN D3 RECEPTOR, FORM A.
FT CHAIN 15 451 VITAMIN D3 RECEPTOR, FORM B.
FT INIT_MET 15 15 FOR FORM B.
FT DNA_BIND 47 112 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 47 67 CA TYPE.
FT ZN_FING 84 107 CA TYPE.
FT DOMAIN 114 215 HINGE.
FT DOMAIN 216 451 LEGND-BINDING.
SO SPOUNCE 451 AA; 51259 MW; 207866ACG8B8F59C C6064;

Query Match 45.18; Score 768; DB 1; Length 451;
Best Local Similarity 48.0%; Pred. No. 1054;
Matches 172; Conservative 71; Mismatches 148; Indels 62; Gaps 10.

QY 8 SWNH-----AFAVCHDFESVSPKSGVNDHDFGCGVQVCAQKAGVHNMI 57
DB 7 SWDQDQSSMAVPIPMADMTVANSISLPDQDQIDKRW--PRIGVGGDQATGFHNMT 63
QY 58 GEGCKGFPRFAKRNARLDCPRKAGCEITRTKQFVACSLKCTLSGKKKMIISQFA 117
DB 64 GEGCKGFPRFAKRNARLDCPRKAGCEITRTKQFVACSLKCTLSGKKKMIISQFA 122
QY 118 VEEERALLKKKKKRTGTGTLGVLTEEGKKMIRELMADKMTPTTSHFKNPLF-- 175
DB 124 VQRRRELLKKKKKELAKESLKKK--LSEGGQVLTITLLEAKHTPTTSSDNRKPRVNR 181
QY 176 -----GVLSGRTELPESDQASREFAKWSGVKRDIAIKVSLDQKQPSVW 224
DB 182 SKESRMAFHSVSVSGQSSPSDSNMVTSDAFAFEDMPFQMSNIDLESQDSISM 241
QY 224 NYKPDASQKELFSLIPIMADKSYMKRGIISPAKVSYNPHLPIDQSLIKLAAPL 283
DB 242 NIEPLH-----LPIPLHADIASVIOKVIGAKKIPNPHDLADQVAILKSSATLV 294
QY 284 GQKRNIVFAELGTWECR--LSVGLD--TAGDPODILTEPMIKFMYMLKQIDHEHY 340
DB 295 IMKRSNDFIMEMISWQSSNDEKKVSLVTVQAGISMLTEEEVKKFVQKIKINIEHEH 354
QY 341 VIMQALSLSPDRHGVGLQRRVVGIOQDPAITLKSYTECHNPGQFAHPEFLKIMAMLPF 400
DB 355 VILMACTISPDHGVQDTSIVESYQDRISDITQVYICRKHPPGSKLYAKKICKLADL 414
QY 401 -----ATPLMOELFG 410
DB 415 RSLNEHSKQYKQSLSPQPSHMLPIPLVEVFG 447

RESULT 7
VDR_RAT STANDARD: PRI: 423 AA.
AC PI3053;
DI 01-JAN-1990 (Rel. 13, Created)
DI 01-JAN-1990 (Rel. 13, Last sequence update)
DI 15-JUL-1999 (Rel. 48, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
GN VDR OR NR1H1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId:10116;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE:89071725; PUBMED 2849110;
RA Burmeister J K., Wise R.J., Maida N., DeLuca H.;
RT "Structure and regulation of the rat 1,25 dihydroxyvitamin D3
RT receptor";
RL Proc. Natl. Acad. Sci. U.S.A. 85:9499-9502(1988).
RN 121;
RP SEQUENCE OF 58-423 FROM N.A.

```



```

Db 80 RLKRCVDIGMKRFFIILDEVRKREMIKKRKEEALKUSLRKLSERQOIIATILDA 138
QY 158 QMKTPTITSHKRNFLP--GVLSKCFELP-----ESIQAPSRF--FAAKWS 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 HIKTYOFTVADEHPPFIRADVSTGSSPRTLLSPQSSSSNDLITSLDMKPEFAS 198
QY 201 QVKKDLVSLKVSQJLKGDEGVSWMNKKPPADSOCKEITSLPLHMADMSYMPKCIISPAV 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 TM-DL-----NEGS---DDPSVTLDLSPLSMPLPHADIVSYIQVVISPAWK 242
QY 261 ISYERDLPELQJSLKGAFFELCOLRPTVNAETGMEGRIISYCLEDT--AGAPQO 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 IPCBRLTSDIQVILKSAAGVILKRSNOSFTMDMSWGQSDPYVDL IOWSRAGHTL 302
QY 318 LLLPMLKRYMMLKQJLQHEEYVYMQAISLSPRPQVLOHRRVQDQJQPAITLKS 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 ELLEPLKRYMMLKQJLQHEEYVYMQAISLSPRPQVLOHRRVQDQJQPAITLKS 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 378 ECKRPPHAFHFLKIMAMITP-----ATPMQELPG 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 KCRHPPGSHQYAKMIQKLAIRSLNPHSKQVRSLSVEPENSMLKTLPLEVEG 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
VDR_HUMAN STANDARD: PRT: 424 AA.
AC Q28047;
DT 01-NOV-1997 (Rel. 45, Created)
DT 01-NOV-1997 (Rel. 45, Last sequence update)
DT 15-JUL-1999 (Rel. 48, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
OS Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
KN 111
RP SEQUENCE FROM N.A.
RX MEDLINE:92044797; PubMed:8880453;
RA Nelbets H.L., Rosworth B.T., Reinhardt T.A.;
RT "Nucleotide sequence of the bovine vitamin D3 receptor."
RL J. Dairy Sci. 79:1313-1315(1996).
CC 1-1 FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC 1-1 SUBCELLULAR LOCATION: NUCLEAR.
CC 1-1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC 1-1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
CC NFI SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DB EMBL: D50209; AAR01543.1; -.
DB HSSP: P04372; LHCO.
DB InterPro: IPR000536; Hormone_rec_14.
DB InterPro: IPR01628; Z1-C4.
DB Pfam: PF00104; hormone_rec_1.
DB Pfam: PF00105; Z1-C4; 1.
DB PRINTS: PR00047; STEROLIDHNR.
DB PRINTS: PR00350; VITAMINDNR.
DB SMART: SM00430; R0C1.1.
DB SMART: SM00439; ZNF_C4; 1.
DB PROSITE: PS00041; NUCLEAR_RECEPTOR; 1.
DB Receptor: Transcription regulation; DNA binding; Nuclear protein;
DB Zinc-finger; Phosphorylation.
DB DNA_BIND 21 86 NUCLEAR RECEPTOR-TYPE.

```

```

FT ZN_PING 21 41 C4-TYPE.
FT ZN_PING 57 81 C4-TYPE.
FT DOMAIN 87 188 HINGE.
FT DOMAIN 189 424 LIGAND-BINDING.
SQ SEQUENCE 424 AA: 47957 MW: 536249262630370 CRC64:

Query Match 34.3% Score 750; DB 1; Length 424;
Best Local Similarity 40.3%; Pred. No. 2,70-53.
Matches 166; Conservative 62; Mismatches 146; Indels 48; Gaps 8;

QY 38 POLGRCGKATGYNFNVWTCGCKGFRPAMKRNARLQCPKCACTETRTTROCAG 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18 PRTGCGKATGYNFNVWTCGCKGFRPAMKRNARLQCPKCACTETRTTROCAG 76
QY 98 RLKRCLESQMKKEMIMSEDAVERKALIKKKSRTTQDPLGVGGLTFQRMIRLQA 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77 RLKRCVDIGMKRFFIILDEVRKREMIKKRKEEALKUSLRKLSERQOIIATILDA 135
QY 158 QMKTPTITSHKRNFLP--GVLSKCFELP-----ESIQAPSRF--FAAKWS 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 HIKTYOFTVADEHPPFIRADVSTGSSPRTLLSPQSSSSNDLITSLDMKPEFAS 194
QY 208 ---SLKVSQJLKGDEGVSWMNKKPPADSOCKEITSLPLHMADMSYMPKCIISPAV 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 MOSSPSNLDLSNDSND- --DPSVTLELSQJLMLPHADIVSYIQVVISPAWK 248
QY 265 RDLPELDQJSLKGAFFELCOLRPTVNAETGMEGRIISYCLEDT--AGAPQO 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 RDLISQJLQVILKSAAGVILKRSNOSFTMDMSWGQSDPYVDL IOWSRAGHTL 308
QY 322 PMLKRYMMLKQJLQHEEYVYMQAISLSPRPQVLOHRRVQDQJQPAITLKS 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 PLKRYMMLKQJLQHEEYVYMQAISLSPRPQVLOHRRVQDQJQPAITLKS 368
QY 382 PDPHAFHFLKIMAMITP-----ATPMQELPG 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 PPHSHQYAKMIQKLAIRSLNPHSKQVRSLSVEPENSMLKTLPLEVEG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
VDR_HUMAN STANDARD: PRT: 427 AA.
AC P11473;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE VITAMIN D3 RECEPTOR (VDR) (1,25-DIHYDROXYVITAMIN D3 RECEPTOR).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
KN 111
RP SEQUENCE FROM N.A.
RX MEDLINE:88217687; PubMed:283767;
RA Baker A.R., McDonnell D.P., Hughes M., Orisp T.M., Mantelstam D.L.,
RA Bussler M.R., Pike J.W., Shine J., O'Malley B.W.;
RT "Cloning and expression of full-length cDNA encoding human vitamin D
RT receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 85:3294-3298(1988).
RN 121
RP SEQUENCE FROM N.A.
RX MEDLINE:92579083; PubMed:1324736;
RA Goto H., Chen K.S., Pfandl J.M., Deluca H.F.;
RT "A single receptor identical with that from rat osteoblast cells
RT mediates the action of 1,25-dihydroxyvitamin D3 in Bb.60 cells."
RL Biochem. Biophys. Acta 1132:103-108(1992).
RN 131
RP SEQUENCE FROM N.A.
RX TISSUE-Lens epithelium;
RA Rae J.L., Shepard A.R.;
RT Submitted (Sep-1997) to the EMBL/GenBank/DDBJ databases.
RL

```

[14]  
 REFERENCE FROM N.A.  
 MEDLINE 97455842; PubMed 97455842  
 Miyamoto K., Kesterson R.A., Yamamoto H., Tokumoto Y., Nishiwaki E.,  
 Terama S., Hara Y., Morita K., Yakoda E., Fike J.W.:  
 "A novel mutation of the human vitamin D receptor chromosomal  
 gene and its functional analysis."  
 Endocrinol. 131:1165-1179(1992).  
 [15]  
 REFERENCE OF 24 90 FROM N.A.  
 Tissue specific blood:  
 MEDLINE 97210272; PubMed 97210272  
 Yu X.-P., Mochlysa H., Hestinger P.T., Manolagas S.C.:  
 "Vitamin D receptor expression in human lymphocytes: signal  
 transduction and characterization by western blots and DNA  
 sequencing."  
 J. Biol. Chem. 266:7568-7595(1991).  
 [16]  
 VARIANTS ASP 43 AND CIN 74  
 MEDLINE 96072761; PubMed 96072761  
 Hughes M.R., Malloy P.J., Kirschner D.C., Kesterson R.A., Fike J.W.:  
 "Point mutations in the human vitamin D receptor gene associated with  
 hypocalcemic rickets."  
 Science 242:1702-1705(1998).  
 [17]  
 VARIANTS CIN 45  
 MEDLINE 96162311; PubMed 96162311  
 Yara H., Ozono K., Miyake H., Nagashima K., Kurama T., Fike J.W.:  
 "A new point mutation in the deoxythymine acid-binding domain of  
 the vitamin D receptor in a kindred with hereditary  
 1,25-dihydroxyvitamin D resistant rickets."  
 J. Clin. Endocrinol. Metab. 76:509-512(1993).  
 [18]  
 VARIANTS CIN 50  
 MEDLINE 9655084; PubMed 9655084  
 Saito T., Ito M., Takada E., Kubohara H., Aihara M., Naito E., Yokota T.,  
 Sone T., Fike J.W., Kurada Y.:  
 "A unique mutation in the vitamin D receptor gene in three Japanese  
 patients with vitamin D-dependent rickets type II: utility of single  
 strand conformation polymorphism analysis for heterozygous carrier  
 detection."  
 Am. J. Hum. Genet. 49:668-674(1991).  
 [19]  
 VARIANTS CIN 80  
 MEDLINE 91125370; PubMed 91125370  
 Sone T., Marx S.J., Itohman U.A., Fike J.W.:  
 "A unique point mutation in the human vitamin D receptor chromosomal  
 gene confers hereditary resistance to 1,25-dihydroxyvitamin D."  
 Mol. Endocrinol. 4:624-631(1990).  
 [20]  
 VARIANTS CIN 81  
 MEDLINE 96194129; PubMed 96194129  
 Malloy P.J., Weisman Y., Feldman D.:  
 "Hereditary 1,25-dihydroxyvitamin D resistant rickets resulting  
 from a mutation in the vitamin D receptor deoxythymine  
 acid binding domain."  
 J. Clin. Endocrinol. Metab. 78:411-416(1994).  
 [21]  
 VARIANTS LBN 274  
 MEDLINE 96415644; PubMed 96415644  
 Kristjansson K., Pott A.R., Howison M., O'Riordan J.L.H., Hughes M.R.:  
 "Two mutations in the hormone binding domain of the vitamin D  
 receptor cause tissue resistance to 1,25-dihydroxyvitamin D."  
 J. Clin. Invest. 92:1216(1993).  
 [22]  
 VARIANTS CIN 45 AND LBN 47  
 MEDLINE 97197677; PubMed 97197677  
 Pott A.R., Howison M., Kristjansson K., Jansz B., Hughes M.R.:  
 "Two mutations causing vitamin D resistant rickets: mapping on the  
 basis of altered hormone receptor DNA binding domain crystal  
 structures."

J. Clin. Endocrinol. (1994) 41:561-566(1994).  
 [23]  
 VARIANTS ASP 45  
 MEDLINE 96272879; PubMed 96272879  
 Liu D.-J., Malloy P.J., Sakai N., Al Ashwal A., Feldman D.:  
 "A novel mutation in the deoxythymine acid-binding domain of the  
 vitamin D receptor causes hereditary 1,25-dihydroxyvitamin D-resistant  
 rickets."  
 J. Clin. Endocrinol. Metab. 81:2564-2569(1999).  
 [24]  
 VARIANTS SER 474 AND CYS 491  
 MEDLINE 97120600; PubMed 97120600  
 Whitfield G.K., Szulczer S.H., Brussler C.A., Bishop J.C.,  
 Gullfash M.A., Jurekka P.W., Thompson P.J., Lee S.M., Zerkow J.E.,  
 Brussler M.R.:  
 "Vitamin D receptors from patients with resistance to 1,25-  
 dihydroxyvitamin D(3): point mutations confer reduced transactivation  
 in response to ligand and impaired interaction with the retinoid X  
 receptor heterodimeric partner."  
 Mol. Endocrinol. 10:1617-1631(1996).  
 [25]  
 VARIANTS CIN 405  
 MEDLINE 97198638; PubMed 97198638  
 Malloy P.J., Kirschner D.C., Gross C., van Muldergom L., Bouillon R.,  
 Feldman D.:  
 "Hereditary vitamin D resistant rickets caused by a novel mutation in  
 the vitamin D receptor that results in decreased affinity for hormone  
 and cellular responsiveness."  
 J. Clin. Invest. 99:297-304(1997).  
 [26]  
 FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D BY CONTROLLING  
 THE EXPRESSION OF HORMONE SENSITIVE GENES.  
 [27]  
 SUBUNIT: INTERACTS WITH SMAD3.  
 [28]  
 SUBCELLULAR LOCATION: NUCLEAR.  
 [29]  
 DOMAIN: COMPOSED OF THREE DOMAINS: A MEDULANTIN N TERMINAL DOMAIN,  
 A DNA-BINDING DOMAIN AND A C TERMINAL SERPINE-BINDING DOMAIN.  
 [30]  
 DISEASE: DEFECTS IN VDR ARE THE CAUSE OF TYPE IIA RICKETS (ALSO  
 KNOWN AS HYPOALCALCAEMIC VITAMIN D RESISTANT RICKETS (HVRD)). HVRD  
 IS MOST FREQUENTLY AN AUTOSOMAL RECESSIVE DISORDER CHARACTERIZED  
 BY SEVERE RICKETS, HYPOCALCAEMIA AND SECONDARY HYPERPARATHYROIDISM.  
 [31]  
 SIMILARITY: RELATIVES TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
 [32]  
 REI SUBFAMILY.  
 [33]  
 CAUTION: IT IS UNCERTAIN WHETHER MET 1 OR MET 4 IS THE INITIATOR.  
 [34]  
 THIS SWISS PROTEIN IS COPYRIGHTED. IT IS PROVIDED THROUGH A COLLABORATION  
 BETWEEN THE SWISS INSTITUTE OF BIOFORMATICS AND THE EMBL OUTSTATION  
 IN THE EUROPEAN BIOFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS  
 USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
 MODIFIED AND THIS STATEMENT IS NOT REMOVED, USURPED BY AND/OR REPRODUCED  
 OUTLIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb.scrib/announc/>  
 or send an email to [license@isb.scrib](mailto:license@isb.scrib)).  
 [35]  
 EMBL: 304258; AAA61273.1  
 EMBL: X67492; CAA47824.1  
 EMBL: AF026260; AAB95155.1  
 EMBL: AF002168; BAA8489.1  
 EMBL: AF002161; BAA8489.1  
 EMBL: AF002162; BAA8489.1  
 EMBL: AF002163; BAA8489.1  
 EMBL: AF002164; BAA8489.1  
 EMBL: AF002165; BAA8489.1  
 EMBL: AF002166; BAA8489.1  
 EMBL: AF002167; BAA8489.1  
 EMBL: AF002168; BAA8489.1  
 EMBL: AF002169; BAA8489.1  
 EMBL: AF002170; BAA8489.1  
 EMBL: AF002171; BAA8489.1  
 EMBL: AF002172; BAA8489.1  
 EMBL: AF002173; BAA8489.1  
 EMBL: AF002174; BAA8489.1  
 EMBL: AF002175; BAA8489.1  
 EMBL: AF002176; BAA8489.1  
 EMBL: AF002177; BAA8489.1  
 EMBL: AF002178; BAA8489.1  
 EMBL: AF002179; BAA8489.1  
 EMBL: AF002180; BAA8489.1  
 EMBL: AF002181; BAA8489.1  
 EMBL: AF002182; BAA8489.1  
 EMBL: AF002183; BAA8489.1  
 EMBL: AF002184; BAA8489.1  
 EMBL: AF002185; BAA8489.1  
 EMBL: AF002186; BAA8489.1  
 EMBL: AF002187; BAA8489.1  
 EMBL: AF002188; BAA8489.1  
 EMBL: AF002189; BAA8489.1  
 EMBL: AF002190; BAA8489.1  
 EMBL: AF002191; BAA8489.1  
 EMBL: AF002192; BAA8489.1  
 EMBL: AF002193; BAA8489.1  
 EMBL: AF002194; BAA8489.1  
 EMBL: AF002195; BAA8489.1  
 EMBL: AF002196; BAA8489.1  
 EMBL: AF002197; BAA8489.1  
 EMBL: AF002198; BAA8489.1  
 EMBL: AF002199; BAA8489.1  
 EMBL: AF002200; BAA8489.1  
 EMBL: AF002201; BAA8489.1  
 EMBL: AF002202; BAA8489.1  
 EMBL: AF002203; BAA8489.1  
 EMBL: AF002204; BAA8489.1  
 EMBL: AF002205; BAA8489.1  
 EMBL: AF002206; BAA8489.1  
 EMBL: AF002207; BAA8489.1  
 EMBL: AF002208; BAA8489.1  
 EMBL: AF002209; BAA8489.1  
 EMBL: AF002210; BAA8489.1  
 EMBL: AF002211; BAA8489.1  
 EMBL: AF002212; BAA8489.1  
 EMBL: AF002213; BAA8489.1  
 EMBL: AF002214; BAA8489.1  
 EMBL: AF002215; BAA8489.1  
 EMBL: AF002216; BAA8489.1  
 EMBL: AF002217; BAA8489.1  
 EMBL: AF002218; BAA8489.1  
 EMBL: AF002219; BAA8489.1  
 EMBL: AF002220; BAA8489.1  
 EMBL: AF002221; BAA8489.1  
 EMBL: AF002222; BAA8489.1  
 EMBL: AF002223; BAA8489.1  
 EMBL: AF002224; BAA8489.1  
 EMBL: AF002225; BAA8489.1  
 EMBL: AF002226; BAA8489.1  
 EMBL: AF002227; BAA8489.1  
 EMBL: AF002228; BAA8489.1  
 EMBL: AF002229; BAA8489.1  
 EMBL: AF002230; BAA8489.1  
 EMBL: AF002231; BAA8489.1  
 EMBL: AF002232; BAA8489.1  
 EMBL: AF002233; BAA8489.1  
 EMBL: AF002234; BAA8489.1  
 EMBL: AF002235; BAA8489.1  
 EMBL: AF002236; BAA8489.1  
 EMBL: AF002237; BAA8489.1  
 EMBL: AF002238; BAA8489.1  
 EMBL: AF002239; BAA8489.1  
 EMBL: AF002240; BAA8489.1  
 EMBL: AF002241; BAA8489.1  
 EMBL: AF002242; BAA8489.1  
 EMBL: AF002243; BAA8489.1  
 EMBL: AF002244; BAA8489.1  
 EMBL: AF002245; BAA8489.1  
 EMBL: AF002246; BAA8489.1  
 EMBL: AF002247; BAA8489.1  
 EMBL: AF002248; BAA8489.1  
 EMBL: AF002249; BAA8489.1  
 EMBL: AF002250; BAA8489.1  
 EMBL: AF002251; BAA8489.1  
 EMBL: AF002252; BAA8489.1  
 EMBL: AF002253; BAA8489.1  
 EMBL: AF002254; BAA8489.1  
 EMBL: AF002255; BAA8489.1  
 EMBL: AF002256; BAA8489.1  
 EMBL: AF002257; BAA8489.1  
 EMBL: AF002258; BAA8489.1  
 EMBL: AF002259; BAA8489.1  
 EMBL: AF002260; BAA8489.1  
 EMBL: AF002261; BAA8489.1  
 EMBL: AF002262; BAA8489.1  
 EMBL: AF002263; BAA8489.1  
 EMBL: AF002264; BAA8489.1  
 EMBL: AF002265; BAA8489.1  
 EMBL: AF002266; BAA8489.1  
 EMBL: AF002267; BAA8489.1  
 EMBL: AF002268; BAA8489.1  
 EMBL: AF002269; BAA8489.1  
 EMBL: AF002270; BAA8489.1  
 EMBL: AF002271; BAA8489.1  
 EMBL: AF002272; BAA8489.1  
 EMBL: AF002273; BAA8489.1  
 EMBL: AF002274; BAA8489.1  
 EMBL: AF002275; BAA8489.1  
 EMBL: AF002276; BAA8489.1  
 EMBL: AF002277; BAA8489.1  
 EMBL: AF002278; BAA8489.1  
 EMBL: AF002279; BAA8489.1  
 EMBL: AF002280; BAA8489.1  
 EMBL: AF002281; BAA8489.1  
 EMBL: AF002282; BAA8489.1  
 EMBL: AF002283; BAA8489.1  
 EMBL: AF002284; BAA8489.1  
 EMBL: AF002285; BAA8489.1  
 EMBL: AF002286; BAA8489.1  
 EMBL: AF002287; BAA8489.1  
 EMBL: AF002288; BAA8489.1  
 EMBL: AF002289; BAA8489.1  
 EMBL: AF002290; BAA8489.1  
 EMBL: AF002291; BAA8489.1  
 EMBL: AF002292; BAA8489.1  
 EMBL: AF002293; BAA8489.1  
 EMBL: AF002294; BAA8489.1  
 EMBL: AF002295; BAA8489.1  
 EMBL: AF002296; BAA8489.1  
 EMBL: AF002297; BAA8489.1  
 EMBL: AF002298; BAA8489.1  
 EMBL: AF002299; BAA8489.1  
 EMBL: AF002300; BAA8489.1  
 EMBL: AF002301; BAA8489.1  
 EMBL: AF002302; BAA8489.1  
 EMBL: AF002303; BAA8489.1  
 EMBL: AF002304; BAA8489.1  
 EMBL: AF002305; BAA8489.1  
 EMBL: AF002306; BAA8489.1  
 EMBL: AF002307; BAA8489.1  
 EMBL: AF002308; BAA8489.1  
 EMBL: AF002309; BAA8489.1  
 EMBL: AF002310; BAA8489.1  
 EMBL: AF002311; BAA8489.1  
 EMBL: AF002312; BAA8489.1  
 EMBL: AF002313; BAA8489.1  
 EMBL: AF002314; BAA8489.1  
 EMBL: AF002315; BAA8489.1  
 EMBL: AF002316; BAA8489.1  
 EMBL: AF002317; BAA8489.1  
 EMBL: AF002318; BAA8489.1  
 EMBL: AF002319; BAA8489.1  
 EMBL: AF002320; BAA8489.1  
 EMBL: AF002321; BAA8489.1  
 EMBL: AF002322; BAA8489.1  
 EMBL: AF002323; BAA8489.1  
 EMBL: AF002324; BAA8489.1  
 EMBL: AF002325; BAA8489.1  
 EMBL: AF002326; BAA8489.1  
 EMBL: AF002327; BAA8489.1  
 EMBL: AF002328; BAA8489.1  
 EMBL: AF002329; BAA8489.1  
 EMBL: AF002330; BAA8489.1  
 EMBL: AF002331; BAA8489.1  
 EMBL: AF002332; BAA8489.1  
 EMBL: AF002333; BAA8489.1  
 EMBL: AF002334; BAA8489.1  
 EMBL: AF002335; BAA8489.1  
 EMBL: AF002336; BAA8489.1  
 EMBL: AF002337; BAA8489.1  
 EMBL: AF002338; BAA8489.1  
 EMBL: AF002339; BAA8489.1  
 EMBL: AF002340; BAA8489.1  
 EMBL: AF002341; BAA8489.1  
 EMBL: AF002342; BAA8489.1  
 EMBL: AF002343; BAA8489.1  
 EMBL: AF002344; BAA8489.1  
 EMBL: AF002345; BAA8489.1  
 EMBL: AF002346; BAA8489.1  
 EMBL: AF002347; BAA8489.1  
 EMBL: AF002348; BAA8489.1  
 EMBL: AF002349; BAA8489.1  
 EMBL: AF002350; BAA8489.1  
 EMBL: AF002351; BAA8489.1  
 EMBL: AF002352; BAA8489.1  
 EMBL: AF002353; BAA8489.1  
 EMBL: AF002354; BAA8489.1  
 EMBL: AF002355; BAA8489.1  
 EMBL: AF002356; BAA8489.1  
 EMBL: AF002357; BAA8489.1  
 EMBL: AF002358; BAA8489.1  
 EMBL: AF002359; BAA8489.1  
 EMBL: AF002360; BAA8489.1  
 EMBL: AF002361; BAA8489.1  
 EMBL: AF002362; BAA8489.1  
 EMBL: AF002363; BAA8489.1  
 EMBL: AF002364; BAA8489.1  
 EMBL: AF002365; BAA8489.1  
 EMBL: AF002366; BAA8489.1  
 EMBL: AF002367; BAA8489.1  
 EMBL: AF002368; BAA8489.1  
 EMBL: AF002369; BAA8489.1  
 EMBL: AF002370; BAA8489.1  
 EMBL: AF002371; BAA8489.1  
 EMBL: AF002372; BAA8489.1  
 EMBL: AF002373; BAA8489.1  
 EMBL: AF002374; BAA8489.1  
 EMBL: AF002375; BAA8489.1  
 EMBL: AF002376; BAA8489.1  
 EMBL: AF002377; BAA8489.1  
 EMBL: AF002378; BAA8489.1  
 EMBL: AF002379; BAA8489.1  
 EMBL: AF002380; BAA8489.1  
 EMBL: AF002381; BAA8489.1  
 EMBL: AF002382; BAA8489.1  
 EMBL: AF002383; BAA8489.1  
 EMBL: AF002384; BAA8489.1  
 EMBL: AF002385; BAA8489.1  
 EMBL: AF002386; BAA8489.1  
 EMBL: AF002387; BAA8489.1  
 EMBL: AF002388; BAA8489.1  
 EMBL: AF002389; BAA8489.1  
 EMBL: AF002390; BAA8489.1  
 EMBL: AF002391; BAA8489.1  
 EMBL: AF002392; BAA8489.1  
 EMBL: AF002393; BAA8489.1  
 EMBL: AF002394; BAA8489.1  
 EMBL: AF002395; BAA8489.1  
 EMBL: AF002396; BAA8489.1  
 EMBL: AF002397; BAA8489.1  
 EMBL: AF002398; BAA8489.1  
 EMBL: AF002399; BAA8489.1  
 EMBL: AF002400; BAA8489.1  
 EMBL: AF002401; BAA8489.1  
 EMBL: AF002402; BAA8489.1  
 EMBL: AF002403; BAA8489.1  
 EMBL: AF002404; BAA8489.1  
 EMBL: AF002405; BAA8489.1  
 EMBL: AF002406; BAA8489.1  
 EMBL: AF002407; BAA8489.1  
 EMBL: AF002408; BAA8489.1  
 EMBL: AF002409; BAA8489.1  
 EMBL: AF002410; BAA8489.1  
 EMBL: AF002411; BAA8489.1  
 EMBL: AF002412; BAA8489.1  
 EMBL: AF002413; BAA8489.1  
 EMBL: AF002414; BAA8489.1  
 EMBL: AF002415; BAA8489.1  
 EMBL: AF002416; BAA8489.1  
 EMBL: AF002417; BAA8489.1  
 EMBL: AF002418; BAA8489.1  
 EMBL: AF002419; BAA8489.1  
 EMBL: AF002420; BAA8489.1  
 EMBL: AF002421; BAA8489.1  
 EMBL: AF002422; BAA8489.1  
 EMBL: AF002423; BAA8489.1  
 EMBL: AF002424; BAA8489.1  
 EMBL: AF002425; BAA8489.1  
 EMBL: AF002426; BAA8489.1  
 EMBL: AF002427; BAA8489.1  
 EMBL: AF002428; BAA8489.1  
 EMBL: AF002429; BAA8489.1  
 EMBL: AF002430; BAA8489.1  
 EMBL: AF002431; BAA8489.1  
 EMBL: AF002432; BAA8489.1  
 EMBL: AF002433; BAA8489.1  
 EMBL: AF002434; BAA8489.1  
 EMBL: AF002435; BAA8489.1  
 EMBL: AF002436; BAA8489.1  
 EMBL: AF002437; BAA8489.1  
 EMBL: AF002438; BAA8489.1  
 EMBL: AF002439; BAA8489.1  
 EMBL: AF002440; BAA8489.1  
 EMBL: AF002441; BAA8489.1  
 EMBL: AF002442; BAA8489.1  
 EMBL: AF002443; BAA8489.1  
 EMBL: AF002444; BAA8489.1  
 EMBL: AF002445; BAA8489.1  
 EMBL: AF002446; BAA8489.1  
 EMBL: AF002447; BAA8489.1  
 EMBL: AF002448; BAA8489.1  
 EMBL: AF002449; BAA8489.1  
 EMBL: AF002450; BAA8489.1  
 EMBL: AF002451; BAA8489.1  
 EMBL: AF002452; BAA8489.1  
 EMBL: AF002453; BAA8489.1  
 EMBL: AF002454; BAA8489.1  
 EMBL: AF002455; BAA8489.1  
 EMBL: AF002456; BAA8489.1  
 EMBL: AF002457; BAA8489.1  
 EMBL: AF002458; BAA8489.1  
 EMBL: AF002459; BAA8489.1  
 EMBL: AF002460; BAA8489.1  
 EMBL: AF002461; BAA8489.1  
 EMBL: AF002462; BAA8489.1  
 EMBL: AF002463; BAA8489.1  
 EMBL: AF002464; BAA8489.1  
 EMBL: AF002465; BAA8489.1  
 EMBL: AF002466; BAA8489.1  
 EMBL: AF002467; BAA8489.1  
 EMBL: AF002468; BAA8489.1  
 EMBL: AF002469; BAA8489.1  
 EMBL: AF002470; BAA8489.1  
 EMBL: AF002471; BAA8489.1  
 EMBL: AF002472; BAA8489.1  
 EMBL: AF002473; BAA8489.1  
 EMBL: AF002474; BAA8489.1  
 EMBL: AF002475; BAA8489.1  
 EMBL: AF002476; BAA8489.1  
 EMBL: AF002477; BAA8489.1  
 EMBL: AF002478; BAA8489.1  
 EMBL: AF002479; BAA8489.1  
 EMBL: AF002480; BAA8489.1  
 EMBL: AF002481; BAA8489.1  
 EMBL: AF002482; BAA8489.1  
 EMBL: AF002483; BAA8489.1  
 EMBL: AF002484; BAA8489.1  
 EMBL: AF002485; BAA8489.1  
 EMBL: AF002486; BAA8489.1  
 EMBL: AF002487; BAA8489.1  
 EMBL: AF002488; BAA8489.1  
 EMBL: AF002489; BAA8489.1  
 EMBL: AF002490; BAA8489.1  
 EMBL: AF002491; BAA8489.1  
 EMBL: AF002492; BAA8489.1  
 EMBL: AF002493; BAA8489.1  
 EMBL: AF002494; BAA8489.1  
 EMBL: AF002495; BAA8489.1  
 EMBL: AF002496; BAA8489.1  
 EMBL: AF002497; BAA8489.1  
 EMBL: AF002498; BAA8489.1  
 EMBL: AF002499; BAA8489.1  
 EMBL: AF002500; BAA8489.1  
 EMBL: AF002501; BAA8489.1  
 EMBL: AF002502; BAA8489.1  
 EMBL: AF002503; BAA8489.1  
 EMBL: AF002504; BAA8489.1  
 EMBL: AF002505; BAA8489.1  
 EMBL: AF002506; BAA8489.1  
 EMBL: AF002507; BAA8489.1  
 EMBL: AF002508; BAA8489.1  
 EMBL: AF002509; BAA8489.1  
 EMBL: AF002510; BAA8489.1  
 EMBL: AF002511; BAA8489.1  
 EMBL: AF002512; BAA8489.1  
 EMBL: AF002513; BAA8489.1  
 EMBL: AF002514; BAA8489.1  
 EMBL: AF002515; BAA8489.1  
 EMBL: AF002516; BAA8489.1  
 EMBL: AF002517; BAA8489.1  
 EMBL: AF002518; BAA8489.1  
 EMBL: AF002519; BAA8489.1  
 EMBL: AF002520; BAA8489.1  
 EMBL: AF002521; BAA8489.1  
 EMBL: AF002522; BAA8489.1  
 EMBL: AF002523; BAA8489.1  
 EMBL: AF002524; BAA8489.1  
 EMBL: AF002525; BAA8489.1  
 EMBL: AF002526; BAA8489.1  
 EMBL: AF002527; BAA8489.1  
 EMBL: AF002528; BAA8489.1  
 EMBL: AF002529; BAA8489.1  
 EMBL: AF002530; BAA8489.1  
 EMBL: AF002531; BAA8489.1  
 EMBL: AF002532; BAA8489.1  
 EMBL: AF002533; BAA8489.1  
 EMBL: AF002534; BAA8489.1  
 EMBL: AF002535; BAA8489.1  
 EMBL: AF002536; BAA8489.1  
 EMBL: AF002537; BAA8489.1  
 EMBL: AF002538; BAA8489.1  
 EMBL: AF002539; BAA8489.1  
 EMBL: AF002540; BAA8489.1  
 EMBL: AF002541; BAA8489.1  
 EMBL: AF002542; BAA8489.1  
 EMBL: AF002543; BAA8489.1  
 EMBL: AF002544; BAA8489.1  
 EMBL: AF002545; BAA8489.1  
 EMBL: AF002546; BAA8489.1  
 EMBL: AF002547; BAA8489.1  
 EMBL: AF002548; BAA8489.1  
 EMBL: AF002549; BAA8489.1  
 EMBL: AF002550; BAA8489.1  
 EMBL: AF002551; BAA8489.1  
 EMBL: AF002552; BAA8489.1  
 EMBL: AF002553; BAA8489.1  
 EMBL: AF002554; BAA8489.1  
 EMBL: AF002555; BAA8489.1  
 EMBL: AF002556; BAA8489.1  
 EMBL: AF002557; BAA8489.1  
 EMBL: AF002558; BAA8489.1  
 EMBL: AF002559; BAA8489.1  
 EMBL: AF002560; BAA8489.1  
 EMBL: AF002561; BAA8489









GenCore version 4.5  
Copyright (c) 1994 - 2000 CompuGen Ltd.

EM protein - protein search, using sw model

Run on: March 2, 2002, 02:36:10 ; Search time 56.22 seconds

(with all alignments)  
1077,139 Million cell updates/sec

Title: 08-09-276-935b-14  
Perfect score: 2187  
Sequence: 1 LEVERKUSENAIRVHCEH.....AMLEFATPLMPLFGTGS 414

Scoring table:  
Gapop 10.0, Gapext 0.5

Searched: 473505 seqs, 146272429 residues

Total number of hits satisfying chosen parameters: 473505

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post-processing: Minimum Match 0.8  
Maximum Match 1008  
1387164 Total 45 summaries

Database:

- 1: SP archaea:\*
- 2: SP bacteria:\*
- 3: SP fungi:\*
- 4: SP human:\*
- 5: SP invertebrate:\*
- 6: SP mammal:\*
- 7: SP mhc:\*
- 8: SP oranelle:\*
- 9: SP phase:\*
- 10: SP plant:\*
- 11: SP rodent:\*
- 12: SP virus:\*
- 13: SP vertebrate:\*
- 14: SP unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 2167   | 99.1        | 473    | 4     | Q9UMN4      |
| 2          | 2158   | 98.7        | 457    | 4     | Q9UJ26      |
| 3          | 2156   | 98.6        | 434    | 4     | Q9UJ27      |
| 4          | 1938.5 | 88.6        | 420    | 4     | Q9UJ23      |
| 5          | 1936.5 | 88.5        | 397    | 4     | Q9UJ24      |
| 6          | 1861   | 84.6        | 479    | 4     | Q9UJ25      |
| 7          | 1700.5 | 77.8        | 411    | 6     | Q9UJ02      |
| 8          | 1631.5 | 74.6        | 442    | 4     | Q9UJ22      |
| 9          | 922.5  | 42.2        | 385    | 13    | Q9U139      |
| 10         | 922.5  | 42.2        | 388    | 13    | Q9UJ24      |
| 11         | 876.5  | 40.1        | 491    | 13    | Q9UJH3      |
| 12         | 771    | 35.3        | 420    | 13    | Q9UJH3      |
| 13         | 748.5  | 34.2        | 425    | 13    | Q9UJH3      |
| 14         | 746    | 34.1        | 453    | 13    | Q9UJH2      |
| 15         | 536.5  | 24.5        | 124    | 11    | Q9UJH2      |
| 16         | 453    | 20.7        | 518    | 5     | Q9UJH2      |
| 17         | 434.5  | 19.9        | 784    | 5     | Q9UJH1      |
| 18         | 431    | 19.7        | 673    | 5     | Q9UJH1      |
| 19         | 428.5  | 19.6        | 680    | 5     | Q9UJH4      |

|    |       |      |     |    |        |
|----|-------|------|-----|----|--------|
| 20 | 424.5 | 19.4 | 472 | 4  | Q9UJH3 |
| 21 | 413.5 | 18.9 | 881 | 5  | Q9UJH8 |
| 22 | 411   | 18.8 | 469 | 11 | Q9UJH3 |
| 23 | 410.5 | 18.6 | 541 | 5  | Q9UJH5 |
| 24 | 406.5 | 18.6 | 484 | 11 | Q9UJH1 |
| 25 | 402.5 | 18.4 | 491 | 5  | Q9UJH5 |
| 26 | 401   | 18.3 | 454 | 13 | Q9UJH2 |
| 27 | 393   | 18.0 | 447 | 13 | Q9UJH3 |
| 28 | 390.5 | 17.9 | 447 | 13 | Q9UJH2 |
| 29 | 387.5 | 17.7 | 444 | 5  | Q9UJH3 |
| 30 | 387   | 17.7 | 560 | 5  | Q9UJH3 |
| 31 | 387   | 17.7 | 570 | 5  | Q9UJH3 |
| 32 | 386   | 17.6 | 444 | 13 | Q9UJH1 |
| 33 | 386   | 17.6 | 444 | 13 | Q9UJH1 |
| 34 | 385.5 | 17.6 | 513 | 5  | Q9UJH3 |
| 35 | 385.5 | 17.6 | 541 | 5  | Q9UJH3 |
| 36 | 384   | 17.6 | 455 | 13 | Q9UJH3 |
| 37 | 383   | 17.5 | 459 | 11 | Q9UJH1 |
| 38 | 382.5 | 17.5 | 452 | 13 | Q9UJH5 |
| 39 | 381   | 17.4 | 367 | 13 | Q9UJH3 |
| 40 | 381   | 17.4 | 448 | 13 | Q9UJH3 |
| 41 | 381   | 17.4 | 555 | 12 | Q9UJH4 |
| 42 | 378   | 17.3 | 456 | 12 | Q9UJH1 |
| 43 | 378   | 17.3 | 555 | 12 | Q9UJH1 |
| 44 | 378   | 17.3 | 582 | 12 | Q9UJH4 |
| 45 | 375   | 17.1 | 453 | 13 | Q9UJH7 |

## ALIGNMENTS

| RESULT | ID   | Q9UMN4       | PRELIMINARY            | PR | 473 AA |
|--------|--|--------------|------------------------|----|--------|
| AC     | Q9UMN4   |              |                        |    |        |
| DT     | 01-MAY-2000  | CTEMBL:01-13 | Created                |    |        |
| DT     | 01-MAY-2000  | CTEMBL:01-13 | Last sequence update   |    |        |
| DI     | 01-JUN-2001  | CTEMBL:01-17 | Last annotation update |    |        |
| DE     | ORPHAN NUCLEAR RECEPTOR  |              |                        |    |        |
| GN     | PAR2   |              |                        |    |        |
| OS     | Homo sapiens (Human)   |              |                        |    |        |
| OC     | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi   |              |                        |    |        |
| OC     | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo  |              |                        |    |        |
| OX     | NCBI_TaxID:9606  |              |                        |    |        |
| RN     | [1]  |              |                        |    |        |
| RP     | SEQUENCE FROM N.A.   |              |                        |    |        |
| RC     | ILISSUE-11576  |              |                        |    |        |
| RA     | MEDLINE:98445550; PubMed:9770465   |              |                        |    |        |
| RA     | Bertilsson G., Heidrich J., Svensson K., Asman M., Jendelewa L., Sydow-Backman M., Ohlsson R., Postlund H., Blomquist P., Beckstrom A. |              |                        |    |        |
| RA     | Identification of a human nuclear receptor defines a new signaling pathway for CYP1A induction.  |              |                        |    |        |
| RT     | Proc. Natl. Acad. Sci. U.S.A. 95:12208-12214 (1998)  |              |                        |    |        |
| RL     | IDENTIFICATION OF A HUMAN NUCLEAR RECEPTOR DEFINES A NEW SIGNALING PATHWAY FOR CYP1A INDUCTION   |              |                        |    |        |
| CC     | 1- SIMILARITY TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY   |              |                        |    |        |
| DR     | EMBL: AF084644; AAC64557.1   |              |                        |    |        |
| DR     | HSSP: P10826; IIRK   |              |                        |    |        |
| DR     | InterPro: IPR000536; Hormone-rec-134   |              |                        |    |        |
| DR     | InterPro: IPR001723; Steroidhormone_recepto  |              |                        |    |        |
| DR     | InterPro: IPR001628; ZF-C4   |              |                        |    |        |
| DR     | Pfam: PF00104; hormone-rec_1   |              |                        |    |        |
| DR     | Pfam: PF00105; ZF-C4_1   |              |                        |    |        |
| DR     | PRINTS: PR00047; STERIODRENER  |              |                        |    |        |
| DR     | SMART: SM00307; HELL_1   |              |                        |    |        |
| DR     | SMART: SM00399; ZNF_C4_1   |              |                        |    |        |
| DR     | PROSITE: PS00031; NUCLEAR_RECEPTOR_1   |              |                        |    |        |
| KW     | DNA-binding; Nuclear protein; Receptor; Transcription regulation; Zinc-finger  |              |                        |    |        |
| SQ     | SEQUENCE 473 AA: 54899 MW: 10490269880016 GRAVY:   |              |                        |    |        |



DR PRINTS: PR000498: STEROIDHORMONE.  
 DR PRINTS: PR000447: STEROIDHORMONE.  
 DR SMART: SM00440: H011: 1.  
 DR SMART: SM00399: Z0F\_C4: 1.  
 DR PROSITE: PS00031: NUCLEAR\_RECEPTOR\_1.  
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation.  
 DR Zinc-finger.  
 DR SOURCE: 434 AA: 49777 MW: AB360823: CAC4200: CIRC04:

Query Match 98.98; Score 2156; DB 4; Length 434;  
 Best Local Similarity 94.78; Pred No. 4.8e-178;  
 Matches 411; Conservative 1; Mismatches 2; Indels 20; Gaps 1;

QY 1 LEVPRKSNMHALFVHCEDTESVPGKPSVNADEVGSPQICRVCGRKATGHEFNVMTCG 60  
 DB 1 MEVPRKSNMHALFVHCEDTESVPGKPSVNADEVGSPQICRVCGRKATGHEFNVMTCG 60  
 QY 61 CKGFFRRAMRNARLKCPKRGACETTRKTRGQACRLKRCLESCKKEMMSDEAVEE 120  
 DB 61 CKGFFRRAMRNARLKCPKRGACETTRKTRGQACRLKRCLESCKKEMMSDEAVEE 120  
 QY 121 KRALIKKKKSEKGTGTPLVQGLTFEGRMMIRLMDQMKTFDTTFSHFNRLPGVLS 180  
 DB 121 KRALIKKKKSEKGTGTPLVQGLTFEGRMMIRLMDQMKTFDTTFSHFNRLPGVLS 180  
 QY 121 KRALIKKKKSEKGTGTPLVQGLTFEGRMMIRLMDQMKTFDTTFSHFNRLPGVLS 180  
 DB 121 KRALIKKKKSEKGTGTPLVQGLTFEGRMMIRLMDQMKTFDTTFSHFNRLPGVLS 180  
 QY 181 GGLPESLQAPSEFAKMSVVKQI/SUKVSLQIRGDSVWYKRPADSGKEIFSTL 240  
 DB 181 GGLPESLQAPSEFAKMSVVKQI/SUKVSLQIRGDSVWYKRPADSGKEIFSTL 240  
 QY 241 PHMAKMSVYMKKQISPAVVISYFRDLPLEQDILSKAAFEICQLFNIVNAETGWE 300  
 DB 241 PHMAKMSVYMKKQISPAVVISYFRDLPLEQDILSKAAFEICQLFNIVNAETGWE 300  
 QY 241 PHMAKMSVYMKKQISPAVVISYFRDLPLEQDILSKAAFEICQLFNIVNAETGWE 300  
 DB 241 PHMAKMSVYMKKQISPAVVISYFRDLPLEQDILSKAAFEICQLFNIVNAETGWE 300  
 QY 401 CGRLSTVEDIAGFPQGLLEPKLKHVMKKQLDHEEYVLMQATISFSDRPGVLAQR 360  
 DB 401 CGRLSTVEDIAGFPQGLLEPKLKHVMKKQLDHEEYVLMQATISFSDRPGVLAQR 360  
 QY 401 CGRLSTVEDIAGFPQGLLEPKLKHVMKKQLDHEEYVLMQATISFSDRPGVLAQR 360  
 DB 401 CGRLSTVEDIAGFPQGLLEPKLKHVMKKQLDHEEYVLMQATISFSDRPGVLAQR 360  
 QY 461 VVDQLQGFATITKSYTECNRPPAHRLFLKIMMLE-----F 400  
 DB 461 VVDQLQGFATITKSYTECNRPPAHRLFLKIMMLE-----F 400  
 QY 401 ATPLMDELFGITGS 434  
 DB 401 ATPLMDELFGITGS 434

RESULT 4  
 Q90J23  
 ID Q90J23 PRELIMINARY; PRT: 420 AA.  
 AC Q90J23;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE NUCLEAR HORMONE RECEPTOR PRR2-C.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE: LIVER.  
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,  
 RA Vissing H.;  
 RA \*Identification of a novel protein isoform of the human nuclear  
 RA hormone receptor PXR/SXR and localization to chromosome 3q12.1  
 RA -13.3.\*  
 RL Eur. J. Hum. Genet. 0:0-0(0).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: TO G4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.  
 DR EMBL: A1009937; CAB55493.1; -  
 DR HSSP: P10826; 1HRA.

DR InterPro: IPR000546; Hormone\_rec\_14.  
 DR InterPro: IPR001723; Steroidhormone\_receptor.  
 DR InterPro: IPR001428; Z1\_C4.  
 DR Pfam: PF00104; hormone\_rec\_1.  
 DR Pfam: PF00105; Z1\_C4\_1.  
 DR PRINTS: PR000498: STEROIDHORMONE.  
 DR PRINTS: PR000447: STEROIDHORMONE.  
 DR SMART: SM00440: H011: 1.  
 DR SMART: SM00399: Z0F\_C4: 1.  
 DR PROSITE: PS00031: NUCLEAR\_RECEPTOR\_1.  
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation.  
 DR Zinc-finger.  
 DR SOURCE: 420 AA: 48149 MW: C949EL25685165: CIRC64;

Query Match 88.68; Score 1938.5; DB 4; Length 420;  
 Best Local Similarity 86.48; Pred No. 2.8e-159;  
 Matches 375; Conservative 0; Mismatches 2; Indels 57; Gaps 2;

QY 1 LEVPRKSNMHALFVHCEDTESVPGKPSVNADEVGSPQICRVCGRKATGHEFNVMTCG 60  
 DB 24 LEVPRKSNMHALFVHCEDTESVPGKPSVNADEVGSPQICRVCGRKATGHEFNVMTCG 64  
 QY 61 CKGFFRRAMRNARLKCPKRGACETTRKTRGQACRLKRCLESCKKEMMSDEAVEE 120  
 DB 61 CKGFFRRAMRNARLKCPKRGACETTRKTRGQACRLKRCLESCKKEMMSDEAVEE 143  
 QY 84 CKGFFRRAMRNARLKCPKRGACETTRKTRGQACRLKRCLESCKKEMMSDEAVEE 143  
 DB 84 CKGFFRRAMRNARLKCPKRGACETTRKTRGQACRLKRCLESCKKEMMSDEAVEE 143  
 QY 121 KRALIKKKKSEKGTGTPLVQGLTFEGRMMIRLMDQMKTFDTTFSHFNRLPGVLS 180  
 DB 144 KRALIKKKKSEKGTGTPLVQGLTFEGRMMIRLMDQMKTFDTTFSHFNRLPGVLS 196  
 QY 181 GGLPESLQAPSEFAKMSVVKQI/SUKVSLQIRGDSVWYKRPADSGKEIFSTL 240  
 DB 197 -----VSQIKRSDSVWYKRPADSGKEIFSTL 226  
 QY 241 PHMAKMSVYMKKQISPAVVISYFRDLPLEQDILSKAAFEICQLFNIVNAETGWE 300  
 DB 227 PHMAKMSVYMKKQISPAVVISYFRDLPLEQDILSKAAFEICQLFNIVNAETGWE 286  
 QY 301 CGRLSTVEDIAGFPQGLLEPKLKHVMKKQLDHEEYVLMQATISFSDRPGVLAQR 360  
 DB 287 CGRLSTVEDIAGFPQGLLEPKLKHVMKKQLDHEEYVLMQATISFSDRPGVLAQR 346  
 QY 361 VVDQLQGFATITKSYTECNRPPAHRLFLKIMMLE-----F 400  
 DB 347 VVDQLQGFATITKSYTECNRPPAHRLFLKIMMLE-----F 406  
 QY 401 ATPLMDELFGITGS 434  
 DB 401 ATPLMDELFGITGS 420

RESULT 5  
 Q90J24  
 ID Q90J24 PRELIMINARY; PRT: 347 AA.  
 AC Q90J24;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE NUCLEAR HORMONE RECEPTOR PRR2-A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE: LIVER.  
 RA Heard D.J., Holloway J., Hansen C., Tommerup N., Aagaard L.,  
 RA Vissing H.;  
 RA \*Identification of a novel protein isoform of the human nuclear  
 RA hormone receptor PXR/SXR and localization to chromosome 3q12.1  
 RA -13.3.\*  
 RL Eur. J. Hum. Genet. 0:0-0(0).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: TO G4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.  
 DR EMBL: A1009937; CAB55493.1; -  
 DR HSSP: P10826; 1HRA.



RA Savas D., Woster M.R., Gillin K.J., Johnson E.F.:  
 RT "The rabbit pregnane X receptor is activated by 11muplin."  
 RL Submitted (SEP-1999) to The EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 20092426; PubMed 10628745;  
 RA Jones S.A., Moore L.R., Sheek J.L., Wisely G.B., Hamilton G.A.,  
 RA McKee D.D., Tomkinson N.C., Lecluyse E.L., Lambert M.H., Willson T.M.,  
 RA Klierer S.A., Moore J.F.:  
 RT "The pregnane X receptor: a promiscuous xenobiotic receptor that has  
 RL diverged during evolution."  
 RL Mol. Endocrinol. 14:27-39(2000).  
 CC -1: SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1: SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.  
 CC EMBL: AF182217; A054426.1;  
 CC EMBL: AF188476; AAF31165.1;  
 CC HSSP: P10826; 1HRA.  
 DR I RefSeq: IP800546; Hormone\_rec\_11q.  
 DR InterPro: IP8001723; Steroidhormone\_receptor.  
 DR InRefPro: IP8001628; Z1-C4.  
 DR Pfam: PF00104; hormone\_rec\_1.  
 DR Pfam: PF00105; Z1-C4; 1.  
 DR PRINTS: PR00598; STEROHORMONR.  
 DR PRINTS: PR00047; STEROIDFINGER.  
 DR SMART: SM00430; HOL1; 1.  
 DR SMART: SM00439; Z0F\_C4; 1.  
 DR PROSITE: PS00041; NUCLEAR\_RECEPTOR; 1.  
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc-finger.  
 SQ SEQUENCE 411 AA; 47309 MW; B8F955EC17EE271A CCK64;

Query Match 77.48; Score 1700.5; DB 6; Length 411;  
 Best Local Similarity 78.78; Pred. No. 9,76-139;  
 Matches 321; Conservative 33; Mismatches 33; Indels 21; Gaps 2;  
 QY 25 GKPVSADDEVGSPQICRGQDKATYHFNVTCECKQFFRKAMKRNARLECPKRGAC 84  
 DB 3 GKPFSADDE:EGPQICRGQDKATYHFNVTCECKQFFRKAMKRNARLECPKRGAC 61  
 QY 85 EITKTRQCGACRLKCGCLSCMKKKKMSDAVERKALLKKKSEFTGPIGVGLT 144  
 DB 62 EITKTRQCGACRLKCGCLSCMKKKKMSDAVERKALLKKKSEFTGPIGVGLT 121  
 QY 145 EEOHMMLEHMDAOMKFDPTFSHFENFKIPVYLSGCTEPESLOAPREKAAWMSQVRK 204  
 DB 122 GEGHMLLEHMDAOMKFDPTFSHFENFKIPVYLSGCTEPESLOAPREKAAWMSQVRK 181  
 QY 205 DIASLKSLSLOKGDGSSVWYKPPASGCKEFTSLPMAWSTYMKGIISEKAVISYF 264  
 DB 182 EIGTKMKLSLOKGDGSSVWYKPPASGCKEFTSLPMAWSTYMKGIISEKAVISYF 241  
 QY 265 RDLPLEDPSILKGAARLFCOLKFNIVNANETGMEGSLSYGCI EDTAGSGFOULLPML 324  
 DB 242 RDLPLEDPSILKGAARLFCOLKFNIVNANETGMEGSLSYGCI EDTAGSGFOULLPML 301  
 QY 425 KKHVKKKQIHEHYVLMQALISPSDKQCVLQHKVVDVLOFOPAITLKSYTFCNRP 384  
 DB 402 KKHVKKKQIHEHYVLMQALISPSDKQCVLQHKVVDVLOFOPAITLKSYTFCNRP 361  
 QY 385 AHRELEKIMAMLE-----FAPIEMOELGII 412  
 DB 462 AHRELEKIMAMLE-----FAPIEMOELGII 409  
 RESULT 8  
 Q90J22 PRELIMINARY; PRI; 342 AA.  
 ID Q90J22;  
 AC Q90J22;  
 DT 01-MAY-2000 (FEMBLrel. 13, created)  
 DT 01-MAY-2000 (FEMBLrel. 13, last sequence update)  
 DT 01-JUN-2001 (FEMBLrel. 17, last annotation update)  
 DE NUCLEAR HORMONE RECEPTOR PRK2-B.

CN PRK2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_Taxid:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-LIVER;  
 RA Heard D.J., Hollway J., Hanson C., Tommstrup N., Aagaard L.,  
 RA Vissing H.:  
 RT "Identification of a novel protein isoform of the human nuclear  
 RT hormone receptor PXR/SXR and localization to chromosome 1q12.1  
 RT -13.3."  
 RL Eur. J. Hum. Genet. 0:0-0(0).  
 CC -1: SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1: SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.  
 CC EMBL: AF009337; CA855494.1;  
 CC HSSP: P10826; 1HRA.  
 DR InterPro: IP800536; Hormone\_rec\_11q.  
 DR InterPro: IP8001723; Steroidhormone\_receptor.  
 DR InRefPro: IP8001628; Z1-C4.  
 DR Pfam: PF00104; hormone\_rec\_1.  
 DR Pfam: PF00105; Z1-C4; 1.  
 DR PRINTS: PR00598; STEROHORMONR.  
 DR PRINTS: PR00047; STEROIDFINGER.  
 DR SMART: SM00430; HOL1; 1.  
 DR SMART: SM00439; Z0F\_C4; 1.  
 DR PROSITE: PS00041; NUCLEAR\_RECEPTOR; 1.  
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc-finger.  
 SQ SEQUENCE 342 AA; 39636 MW; 7C9467C4E900DAVA CCK64;

Query Match 74.68; Score 1631.5; DB 4; Length 442;  
 Best Local Similarity 84.78; Pred. No. 70-185;  
 Matches 321; Conservative 0; Mismatches 1; Indels 57; Gaps 2;  
 QY 56 MTEGCKGFFRRAMKKNARLKCPFKRGACETTRKTRQCGVQRLKRLKESGKKEMMSI 115  
 DB 1 MTEGCKGFFRRAMKKNARLKCPFKRGACETTRKTRQCGVQRLKRLKESGKKEMMSI 60  
 QY 116 FAVFERALLKKKRSSTGTGPIGVGLTTHQKMMIRLMAOMKFDPTFSHFKNRSLP 175  
 DB 61 FAVFERALLKKKRSSTGTGPIGVGLTTHQKMMIRLMAOMKFDPTFSHFKNRSLP 118  
 QY 176 GVLSNGCELPESLOANSREAAWMSQVRKDIASLKSLSLOKGDGSSVWYKPPASGCKE 245  
 DB 119 -----VSIOLKGDGSSVWYKPPASGCKE 144  
 QY 236 IESLLEPMADMSYMKGIISEKAVISYRDLPELQI SLKGAARLFCOLKFNIVNAN 295  
 DB 144 IESLLEPMADMSYMKGIISEKAVISYRDLPELQI SLKGAARLFCOLKFNIVNAN 204  
 QY 296 TGMEGCSLSYGLDAGCHQOULLEPMKKNHMKKQIHEHYVLMQALISPSDKQCVL 355  
 DB 204 TGMEGCSLSYGLDAGCHQOULLEPMKKNHMKKQIHEHYVLMQALISPSDKQCVL 244  
 QY 356 VLQHRVVDVLOFOPAITLKSYTFCNRPVABHLEFKIMAMLE----- 499  
 DB 204 VLQHRVVDVLOFOPAITLKSYTFCNRPVABHLEFKIMAMLE----- 424  
 QY 400 -----FAPIEMOELGII 414  
 DB 424 DIHPPATIMQELPTIGS 342  
 RESULT 9  
 Q91839 PRELIMINARY; PRI; 486 AA.  
 ID Q91839;  
 AC Q91839;  
 DT 01-NOV-1996 (FEMBLrel. 01, created)  
 DT 01-NOV-1996 (FEMBLrel. 01, last sequence update)  
 DT 01-JUN-2001 (FEMBLrel. 17, last annotation update)  
 DE ORPHAN NUCLEAR RECEPTOR OF STEROID/HYPERLIP SUPERFAMILY.

001 Xenopus laevis (African clawed frog).  
 002 Eukaryotic Metazoa: Chordata: Vertebrata: Euteleostomi:  
 003 Amphibia: Batrachia: Anura: Mesobatrachia: Pipridae: Pipidae:  
 004 Xenopus laevis.  
 005 NCBI Jukid B4552.  
 006  
 007  
 008  
 009  
 010  
 011  
 012  
 013  
 014  
 015  
 016  
 017  
 018  
 019  
 020  
 021  
 022  
 023  
 024  
 025  
 026  
 027  
 028  
 029  
 030  
 031  
 032  
 033  
 034  
 035  
 036  
 037  
 038  
 039  
 040  
 041  
 042  
 043  
 044  
 045  
 046  
 047  
 048  
 049  
 050  
 051  
 052  
 053  
 054  
 055  
 056  
 057  
 058  
 059  
 060  
 061  
 062  
 063  
 064  
 065  
 066  
 067  
 068  
 069  
 070  
 071  
 072  
 073  
 074  
 075  
 076  
 077  
 078  
 079  
 080  
 081  
 082  
 083  
 084  
 085  
 086  
 087  
 088  
 089  
 090  
 091  
 092  
 093  
 094  
 095  
 096  
 097  
 098  
 099  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546  
 547  
 548  
 549  
 550  
 551  
 552  
 553  
 554  
 555  
 556  
 557  
 558  
 559  
 560  
 561  
 562  
 563  
 564  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612  
 613  
 614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629  
 630  
 631  
 632  
 633  
 634  
 635  
 636  
 637  
 638  
 639  
 640  
 641  
 642  
 643  
 644  
 645  
 646  
 647  
 648  
 649  
 650  
 651  
 652  
 653  
 654  
 655  
 656  
 657  
 658  
 659  
 660  
 661  
 662  
 663  
 664  
 665  
 666  
 667  
 668  
 669  
 670  
 671  
 672  
 673  
 674  
 675  
 676  
 677  
 678  
 679  
 680  
 681  
 682  
 683  
 684  
 685  
 686  
 687  
 688  
 689  
 690  
 691  
 692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702  
 703  
 704  
 705  
 706  
 707  
 708  
 709  
 710  
 711  
 712  
 713  
 714  
 715  
 716  
 717  
 718  
 719  
 720  
 721  
 722  
 723  
 724  
 725  
 726  
 727  
 728  
 729  
 730  
 731  
 732  
 733  
 734  
 735  
 736  
 737  
 738  
 739  
 740  
 741  
 742  
 743  
 744  
 745  
 746  
 747  
 748  
 749  
 750  
 751  
 752  
 753  
 754  
 755  
 756  
 757  
 758  
 759  
 760  
 761  
 762  
 763  
 764  
 765  
 766  
 767  
 768  
 769  
 770  
 771  
 772  
 773  
 774  
 775  
 776  
 777  
 778  
 779  
 780  
 781  
 782  
 783  
 784  
 785  
 786  
 787  
 788  
 789  
 790  
 791  
 792  
 793  
 794  
 795  
 796  
 797  
 798  
 799  
 800  
 801  
 802  
 803  
 804  
 805  
 806  
 807  
 808  
 809  
 810  
 811  
 812  
 813  
 814  
 815  
 816  
 817  
 818  
 819  
 820  
 821  
 822  
 823  
 824  
 825  
 826  
 827  
 828  
 829  
 830  
 831  
 832  
 833  
 834  
 835  
 836  
 837  
 838  
 839  
 840  
 841  
 842  
 843  
 844  
 845  
 846  
 847  
 848  
 849  
 850  
 851  
 852  
 853  
 854  
 855  
 856  
 857  
 858  
 859  
 860  
 861  
 862  
 863  
 864  
 865  
 866  
 867  
 868  
 869  
 870  
 871  
 872  
 873  
 874  
 875  
 876  
 877  
 878  
 879  
 880  
 881  
 882  
 883  
 884  
 885  
 886  
 887  
 888  
 889  
 890  
 891  
 892  
 893  
 894  
 895  
 896  
 897  
 898  
 899  
 900  
 901  
 902  
 903  
 904  
 905  
 906  
 907  
 908  
 909  
 910  
 911  
 912  
 913  
 914  
 915  
 916  
 917  
 918  
 919  
 920  
 921  
 922  
 923  
 924  
 925  
 926  
 927  
 928  
 929  
 930  
 931  
 932  
 933  
 934  
 935  
 936  
 937  
 938  
 939  
 940  
 941  
 942  
 943  
 944  
 945  
 946  
 947  
 948  
 949  
 950  
 951  
 952  
 953  
 954  
 955  
 956  
 957  
 958  
 959  
 960  
 961  
 962  
 963  
 964  
 965  
 966  
 967  
 968  
 969  
 970  
 971  
 972  
 973  
 974  
 975  
 976  
 977  
 978  
 979  
 980  
 981  
 982  
 983  
 984  
 985  
 986  
 987  
 988  
 989  
 990  
 991  
 992  
 993  
 994  
 995  
 996  
 997  
 998  
 999  
 1000

| Query Match           | 45.38   | Score 771   | 18 13     | Length 420 |
|-----------------------|---|---|-----------|------------|
| Best Local Similarity | 41.08   | Prod. No. 1,967-98  |           |            |
| Matches 170           | Conservative 66   | Mismatches 137  | Indels 42 | Gaps 11    |
| QY                    | 29  | VNADE-EEVSDQIGVCGKATGATGHPNVMIGCGKCPYRANKNANLROPEPKATLTI    | H7        |            |
| DB                    | 11  | VGPDEEFRRNAPRGCGVGKATGATGHPNPMATDEGKPKFKRSKSKASTQPP-NGSGTTI | 69        |            |
| QY                    | 88  | RKTRPGCGATRLKCLESSMKREIMISDEAVERALLIKRKSEKSGTGTLPVVGITTEQ   | 147       |            |
| DB                    | 70  | KQNRHCGACGRLKCTIDTGMKKEPILITDEYQREKEMIKKKEFAARAMPRL-IMPQ    | 148       |            |
| QY                    | 148   | RKMLREIMDQMKPTFTESHEKNEP---GVLSSTELPESHQADEAKKSGVYK         | 204       |            |
| DB                    | 129   | ARMSSIVNEHNTYASISQPSRPVPRVPRVPTKRSKAAISHLSTASSSTNHSP        | 199       |            |
| QY                    | 205   | DLCSEAK---SLQKGRHSSVNNVTPADSSGKELPSLIDPMADSTYMKKLTISAKVI    | 261       |            |
| DB                    | 189   | SV-DTKNPNSSLIMVQDGA---SSPSSEENTKLSMPLADVSYSTQVYGVAKM        | 243       |            |
| QY                    | 262   | SEYSDPILDEQISLKGAFELQIRENTVNAETGWN-SEISYLED IANDEQV         | 318       |            |
| DB                    | 244   | PGPDPAHFDQMLKSSATLIMLRKSNSTLDMKSNQDGPYKQINQVKAHIE           | 303       |            |
| QY                    | 319   | LLEMLKFMMLKRLDDEEVEYLMQALISLSTIRFQVLDHVVNDLDQFATLKSTIE      | 478       |            |
| DB                    | 404   | LLELVKFGVQIKLIMDEEIVLLMGITLSPDRVQDAHVAQVQDIPKAVYIK          | 463       |            |
| QY                    | 479   | CNRQVABRLFLIKIMAMLEF-----                                   | 410       |            |
| DB                    | 364   | IN-HGCGILYAKMLOKLOLAKSI NEHSKUYASTSPQREHSMQITPLVLEVG        | 416       |            |
| RESULI                | 13  |   |           |            |
| Q91B74                |   |   |           |            |
| AC                    | Q91B74  | PROLIMINARY   | PRT       | 425 AA.    |
| DI                    | 01-OCT-2000   | (ITEM1:rel. 15, Created)                                    |           |            |
| DT                    | 01-OCT-2000   | (ITEM1:rel. 15, last sequence update)                       |           |            |
| DI                    | 01-JUN-2001   | (ITEM1:rel. 17, last annotation update)                     |           |            |
| DI                    | VITAMIN D RECEPTOR B.   |   |           |            |
| DI                    | VDRB.   |   |           |            |
| DI                    | Parathyroid hys. olivaceus (Ploveret).                            |   |           |            |
| DI                    | Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: |   |           |            |



01-JUN-2001 (TEMPREL). 17, (created)  
 01-JUN-2001 (TEMPREL). 17, last sequence update)  
 01-JUN-2001 (TEMPREL). 17, last annotation update)  
 DE NUCLEAR RECEPTOR SUBFAMILY 1, GROUP 1, MEMBER 2 (FRAGMENT).  
 CN NR112.  
 OS Mus musculus (mouse).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;  
 CK Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 DX NCBI\_taxonomy:10090;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RC STRAIN: C57BL/6J; TISSUE: EMBRYONIC LIVER;  
 RX MEDLINE: 21085660; PubMed: 11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Atakawa T., Hara A., Fukunishi Y., Komu H., Adachi J., Fukuda S.,  
 RA Aizawa K., Iwata M., Nishii K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito K.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King H., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide T., Pesole G., Quackenbush J.,  
 RA Schreml L.M., Stambli F., Suzuki K., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Botelli D., Boujard N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Holman M., Hume D.A., Kamiya K., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez L., Sakamoto N.,  
 RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,  
 RA Hayashizaki Y.;  
 RT "National annotation of a full-length mouse cDNA collection.";  
 RL Nucleic Acids Res. 29(20):409-415 (2001).  
 CC -1: SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1: SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.  
 DR EMBL: AK011108; BAB27404.1; -;  
 DR MGI: MGI:147040; NF112.  
 DR InterPro: IPR001628; Z1-C4.  
 DR Pfam: PF00105; Z1-C4.1.  
 DR PRINTS: PR00047; STROIDFINGER.  
 DR SMART: SM00399; ZNF345.1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1  
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc finger.  
 FT NEN\_1ER 124 124  
 SEQUENCE 124 AA; 1452 MW; 21F83A34B248B6C1 CMC64;

Query Match 24.58; Score 536.5; DB 11; Length 124;  
 Best local Similarity 78.4%; Pred. No. 7e-39;  
 Matches 98; Conservative 10; Mismatches 16; Indels 1; Gaps 1;  
 YV 3 VPKPSNNHAIHVCHDETSVPRPSVNDPFGVGSGJCRVGGKAFVYHFNVMTCGCK 62  
 DB 1 MKPFSMSKRVLYVTEADSNLEEP-INVREEDGLOICRVGGDKANGYHFNVMTCGCK 59  
 YV 63 GFFRRAKRNARICPPRRGACELIKIRQCQACIKKCLSCSKMKRMNSDFAVPER 122  
 DB 60 GFFRRAKRNARICPPRRGACELIKIRQCQACIKKCLSCSKMKRMNSDFAVPER 119  
 YV 123 ALIKR 127  
 DB 120 ALIKR 124

Search completed: March 2, 2002, 02:44:32  
 Job time: 182 sec













[illegible][illegible]



[illegible]

|     |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |       |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|
| 146 | 1450 | 1470 | 1480 | 1490 | 1500 | 1510 | 1520 | 1530 | 1540 | 1550 | 1560 | 1570 | 1580 | 1590 | 1600 | 1610 | 1620 | 1630 | 1640 | 1650 | 1660 | 1670 | 1680 | 1690 | 1700 | 1710 | 1720 | 1730 | 1740 | 1750 | 1760 | 1770 | 1780 | 1790 | 1800 | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 | 2030 | 2040 | 2050 | 2060 | 2070 | 2080 | 2090 | 2100 | 2110 | 2120 | 2130 | 2140 | 2150 | 2160 | 2170 | 2180 | 2190 | 2200 | 2210 | 2220 | 2230 | 2240 | 2250 | 2260 | 2270 | 2280 | 2290 | 2300 | 2310 | 2320 | 2330 | 2340 | 2350 | 2360 | 2370 | 2380 | 2390 | 2400 | 2410 | 2420 | 2430 | 2440 | 2450 | 2460 | 2470 | 2480 | 2490 | 2500 | 2510 | 2520 | 2530 | 2540 | 2550 | 2560 | 2570 | 2580 | 2590 | 2600 | 2610 | 2620 | 2630 | 2640 | 2650 | 2660 | 2670 | 2680 | 2690 | 2700 | 2710 | 2720 | 2730 | 2740 | 2750 | 2760 | 2770 | 2780 | 2790 | 2800 | 2810 | 2820 | 2830 | 2840 | 2850 | 2860 | 2870 | 2880 | 2890 | 2900 | 2910 | 2920 | 2930 | 2940 | 2950 | 2960 | 2970 | 2980 | 2990 | 3000 | 3010 | 3020 | 3030 | 3040 | 3050 | 3060 | 3070 | 3080 | 3090 | 3100 | 3110 | 3120 | 3130 | 3140 | 3150 | 3160 | 3170 | 3180 | 3190 | 3200 | 3210 | 3220 | 3230 | 3240 | 3250 | 3260 | 3270 | 3280 | 3290 | 3300 | 3310 | 3320 | 3330 | 3340 | 3350 | 3360 | 3370 | 3380 | 3390 | 3400 | 3410 | 3420 | 3430 | 3440 | 3450 | 3460 | 3470 | 3480 | 3490 | 3500 | 3510 | 3520 | 3530 | 3540 | 3550 | 3560 | 3570 | 3580 | 3590 | 3600 | 3610 | 3620 | 3630 | 3640 | 3650 | 3660 | 3670 | 3680 | 3690 | 3700 | 3710 | 3720 | 3730 | 3740 | 3750 | 3760 | 3770 | 3780 | 3790 | 3800 | 3810 | 3820 | 3830 | 3840 | 3850 | 3860 | 3870 | 3880 | 3890 | 3900 | 3910 | 3920 | 3930 | 3940 | 3950 | 3960 | 3970 | 3980 | 3990 | 4000 | 4010 | 4020 | 4030 | 4040 | 4050 | 4060 | 4070 | 4080 | 4090 | 4100 | 4110 | 4120 | 4130 | 4140 | 4150 | 4160 | 4170 | 4180 | 4190 | 4200 | 4210 | 4220 | 4230 | 4240 | 4250 | 4260 | 4270 | 4280 | 4290 | 4300 | 4310 | 4320 | 4330 | 4340 | 4350 | 4360 | 4370 | 4380 | 4390 | 4400 | 4410 | 4420 | 4430 | 4440 | 4450 | 4460 | 4470 | 4480 | 4490 | 4500 | 4510 | 4520 | 4530 | 4540 | 4550 | 4560 | 4570 | 4580 | 4590 | 4600 | 4610 | 4620 | 4630 | 4640 | 4650 | 4660 | 4670 | 4680 | 4690 | 4700 | 4710 | 4720 | 4730 | 4740 | 4750 | 4760 | 4770 | 4780 | 4790 | 4800 | 4810 | 4820 | 4830 | 4840 | 4850 | 4860 | 4870 | 4880 | 4890 | 4900 | 4910 | 4920 | 4930 | 4940 | 4950 | 4960 | 4970 | 4980 | 4990 | 5000 | 5010 | 5020 | 5030 | 5040 | 5050 | 5060 | 5070 | 5080 | 5090 | 5100 | 5110 | 5120 | 5130 | 5140 | 5150 | 5160 | 5170 | 5180 | 5190 | 5200 | 5210 | 5220 | 5230 | 5240 | 5250 | 5260 | 5270 | 5280 | 5290 | 5300 | 5310 | 5320 | 5330 | 5340 | 5350 | 5360 | 5370 | 5380 | 5390 | 5400 | 5410 | 5420 | 5430 | 5440 | 5450 | 5460 | 5470 | 5480 | 5490 | 5500 | 5510 | 5520 | 5530 | 5540 | 5550 | 5560 | 5570 | 5580 | 5590 | 5600 | 5610 | 5620 | 5630 | 5640 | 5650 | 5660 | 5670 | 5680 | 5690 | 5700 | 5710 | 5720 | 5730 | 5740 | 5750 | 5760 | 5770 | 5780 | 5790 | 5800 | 5810 | 5820 | 5830 | 5840 | 5850 | 5860 | 5870 | 5880 | 5890 | 5900 | 5910 | 5920 | 5930 | 5940 | 5950 | 5960 | 5970 | 5980 | 5990 | 6000 | 6010 | 6020 | 6030 | 6040 | 6050 | 6060 | 6070 | 6080 | 6090 | 6100 | 6110 | 6120 | 6130 | 6140 | 6150 | 6160 | 6170 | 6180 | 6190 | 6200 | 6210 | 6220 | 6230 | 6240 | 6250 | 6260 | 6270 | 6280 | 6290 | 6300 | 6310 | 6320 | 6330 | 6340 | 6350 | 6360 | 6370 | 6380 | 6390 | 6400 | 6410 | 6420 | 6430 | 6440 | 6450 | 6460 | 6470 | 6480 | 6490 | 6500 | 6510 | 6520 | 6530 | 6540 | 6550 | 6560 | 6570 | 6580 | 6590 | 6600 | 6610 | 6620 | 6630 | 6640 | 6650 | 6660 | 6670 | 6680 | 6690 | 6700 | 6710 | 6720 | 6730 | 6740 | 6750 | 6760 | 6770 | 6780 | 6790 | 6800 | 6810 | 6820 | 6830 | 6840 | 6850 | 6860 | 6870 | 6880 | 6890 | 6900 | 6910 | 6920 | 6930 | 6940 | 6950 | 6960 | 6970 | 6980 | 6990 | 7000 | 7010 | 7020 | 7030 | 7040 | 7050 | 7060 | 7070 | 7080 | 7090 | 7100 | 7110 | 7120 | 7130 | 7140 | 7150 | 7160 | 7170 | 7180 | 7190 | 7200 | 7210 | 7220 | 7230 | 7240 | 7250 | 7260 | 7270 | 7280 | 7290 | 7300 | 7310 | 7320 | 7330 | 7340 | 7350 | 7360 | 7370 | 7380 | 7390 | 7400 | 7410 | 7420 | 7430 | 7440 | 7450 | 7460 | 7470 | 7480 | 7490 | 7500 | 7510 | 7520 | 7530 | 7540 | 7550 | 7560 | 7570 | 7580 | 7590 | 7600 | 7610 | 7620 | 7630 | 7640 | 7650 | 7660 | 7670 | 7680 | 7690 | 7700 | 7710 | 7720 | 7730 | 7740 | 7750 | 7760 | 7770 | 7780 | 7790 | 7800 | 7810 | 7820 | 7830 | 7840 | 7850 | 7860 | 7870 | 7880 | 7890 | 7900 | 7910 | 7920 | 7930 | 7940 | 7950 | 7960 | 7970 | 7980 | 7990 | 8000 | 8010 | 8020 | 8030 | 8040 | 8050 | 8060 | 8070 | 8080 | 8090 | 8100 | 8110 | 8120 | 8130 | 8140 | 8150 | 8160 | 8170 | 8180 | 8190 | 8200 | 8210 | 8220 | 8230 | 8240 | 8250 | 8260 | 8270 | 8280 | 8290 | 8300 | 8310 | 8320 | 8330 | 8340 | 8350 | 8360 | 8370 | 8380 | 8390 | 8400 | 8410 | 8420 | 8430 | 8440 | 8450 | 8460 | 8470 | 8480 | 8490 | 8500 | 8510 | 8520 | 8530 | 8540 | 8550 | 8560 | 8570 | 8580 | 8590 | 8600 | 8610 | 8620 | 8630 | 8640 | 8650 | 8660 | 8670 | 8680 | 8690 | 8700 | 8710 | 8720 | 8730 | 8740 | 8750 | 8760 | 8770 | 8780 | 8790 | 8800 | 8810 | 8820 | 8830 | 8840 | 8850 | 8860 | 8870 | 8880 | 8890 | 8900 | 8910 | 8920 | 8930 | 8940 | 8950 | 8960 | 8970 | 8980 | 8990 | 9000 | 9010 | 9020 | 9030 | 9040 | 9050 | 9060 | 9070 | 9080 | 9090 | 9100 | 9110 | 9120 | 9130 | 9140 | 9150 | 9160 | 9170 | 9180 | 9190 | 9200 | 9210 | 9220 | 9230 | 9240 | 9250 | 9260 | 9270 | 9280 | 9290 | 9300 | 9310 | 9320 | 9330 | 9340 | 9350 | 9360 | 9370 | 9380 | 9390 | 9400 | 9410 | 9420 | 9430 | 9440 | 9450 | 9460 | 9470 | 9480 | 9490 | 9500 | 9510 | 9520 | 9530 | 9540 | 9550 | 9560 | 9570 | 9580 | 9590 | 9600 | 9610 | 9620 | 9630 | 9640 | 9650 | 9660 | 9670 | 9680 | 9690 | 9700 | 9710 | 9720 | 9730 | 9740 | 9750 | 9760 | 9770 | 9780 | 9790 | 9800 | 9810 | 9820 | 9830 | 9840 | 9850 | 9860 | 9870 | 9880 | 9890 | 9900 | 9910 | 9920 | 9930 | 9940 | 9950 | 9960 | 9970 | 9980 | 9990 | 10000 |
| 146 | 1450 | 1470 | 1480 | 1490 | 1500 | 1510 | 1520 | 1530 | 1540 | 1550 | 1560 | 1570 | 1580 | 1590 | 1600 | 1610 | 1620 | 1630 | 1640 | 1650 | 1660 | 1670 | 1680 | 1690 | 1700 | 1710 | 1720 | 1730 | 1740 | 1750 | 1760 | 1770 | 1780 | 1790 | 1800 | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 | 2030 | 2040 | 2050 | 2060 | 2070 | 2080 | 2090 | 2100 | 2110 | 2120 | 2130 | 2140 | 2150 | 2160 | 2170 | 2180 | 2190 | 2200 | 2210 | 2220 | 2230 | 2240 | 2250 | 2260 | 2270 | 2280 | 2290 | 2300 | 2310 | 2320 | 2330 | 2340 | 2350 | 2360 | 2370 | 2380 | 2390 | 2400 | 2410 | 2420 | 2430 | 2440 | 2450 | 2460 | 2470 | 2480 | 2490 | 2500 | 2510 | 2520 | 2530 | 2540 | 2550 | 2560 | 2570 | 2580 | 2590 | 2600 | 2610 | 2620 | 2630 | 2640 | 2650 | 2660 | 2670 | 2680 | 2690 | 2700 | 2710 | 2720 | 2730 | 2740 | 2750 | 2760 | 2770 | 2780 | 2790 | 2800 | 2810 | 2820 | 2830 | 2840 | 2850 | 2860 | 2870 | 2880 | 2890 | 2900 | 2910 | 2920 | 2930 | 2940 | 2950 | 2960 | 2970 | 2980 | 2990 | 3000 | 3010 | 3020 | 3030 | 3040 | 3050 | 3060 | 3070 | 3080 | 3090 | 3100 | 3110 | 3120 | 3130 | 3140 | 3150 | 3160 | 3170 | 3180 | 3190 | 3200 | 3210 | 3220 | 3230 | 3240 | 3250 | 3260 | 3270 | 3280 | 3290 | 3300 | 3310 | 3320 | 3330 | 3340 | 3350 | 3360 | 3370 | 3380 | 3390 | 3400 | 3410 | 3420 | 3430 | 3440 | 3450 | 3460 | 3470 | 3480 | 3490 | 3500 | 3510 | 3520 | 3530 | 3540 | 3550 | 3560 | 3570 | 3580 | 3590 | 3600 | 3610 | 3620 | 3630 | 3640 | 3650 | 3660 | 3670 | 3680 | 3690 | 3700 | 3710 | 3720 | 3730 | 3740 | 3750 | 3760 | 3770 | 3780 | 3790 | 3800 | 3810 | 3820 | 3830 | 3840 | 3850 | 3860 | 3870 | 3880 | 3890 | 3900 | 3910 | 3920 | 3930 | 3940 | 3950 | 3960 | 3970 | 3980 | 3990 | 4000 | 4010 | 4020 | 4030 | 4040 | 4050 | 4060 | 4070 | 4080 | 4090 | 4100 | 4110 | 4120 | 4130 | 4140 | 4150 | 4160 | 4170 | 4180 | 4190 | 4200 | 4210 | 4220 | 4230 | 4240 | 4250 | 4260 | 4270 | 4280 | 4290 | 4300 | 4310 | 4320 | 4330 | 4340 | 4350 | 4360 | 4370 | 4380 | 4390 | 4400 | 4410 | 4420 | 4430 | 4440 | 4450 | 4460 | 4470 | 4480 | 4490 | 4500 | 4510 | 4520 | 4530 | 4540 | 4550 | 4560 | 4570 | 4580 | 4590 | 4600 | 4610 | 4620 | 4630 | 4640 | 4650 | 4660 | 4670 | 4680 | 4690 | 4700 | 4710 | 4720 | 4730 | 4740 | 4750 | 4760 | 4770 | 4780 | 4790 | 4800 | 4810 | 4820 | 4830 | 4840 | 4850 | 4860 | 4870 | 4880 | 4890 | 4900 | 4910 | 4920 | 4930 | 4940 | 4950 | 4960 | 4970 | 4980 | 4990 | 5000 | 5010 | 5020 | 5030 | 5040 | 5050 | 5060 | 5070 | 5080 | 5090 | 5100 | 5110 | 5120 | 5130 | 5140 | 5150 | 5160 | 5170 | 5180 | 5190 | 5200 | 5210 | 5220 | 5230 | 5240 | 5250 | 5260 | 5270 | 5280 | 5290 | 5300 | 5310 | 5320 | 5330 | 5340 | 5350 | 5360 | 5370 | 5380 | 5390 | 5400 | 5410 | 5420 | 5430 | 5440 | 5450 | 5460 | 5470 | 5480 | 5490 | 5500 | 5510 | 5520 | 5530 | 5540 | 5550 | 5560 | 5570 | 5580 | 5590 | 5600 | 5610 | 5620 | 5630 | 5640 | 5650 | 5660 | 5670 | 5680 | 5690 | 5700 | 5710 | 5720 | 5730 | 5740 | 5750 | 5760 | 5770 | 5780 | 5790 | 5800 | 5810 | 5820 | 5830 | 5840 | 5850 | 5860 | 5870 | 5880 | 5890 | 5900 | 5910 | 5920 | 5930 | 5940 | 5950 | 5960 | 5970 | 5980 | 5990 | 6000 | 6010 | 6020 | 6030 | 6040 | 6050 | 6060 | 6070 | 6080 | 6090 | 6100 | 6110 | 6120 | 6130 | 6140 | 6150 | 6160 | 6170 | 6180 | 6190 | 6200 | 6210 | 6220 | 6230 | 6240 | 6250 | 6260 | 6270 | 6280 | 6290 | 6300 | 6310 | 6320 | 6330 | 6340 | 6350 | 6360 | 6370 | 6380 | 6390 | 6400 | 6410 | 6420 | 6430 | 6440 | 6450 | 6460 | 6470 | 6480 | 6490 | 6500 | 6510 | 6520 | 6530 | 6540 | 6550 | 6560 | 6570 | 6580 | 6590 | 6600 | 6610 | 6620 | 6630 |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |       |

/tissue\_type "Intestine"  
 /tissue\_type "colon"  
 1771..4013  
 /gene "PRK2"  
 1771..4013  
 /cds "PRK2"  
 1771..4013  
 /note "isoform with longer A/h domain. translation  
 initiates at GUG-1771. 5 fold less efficiently than at  
 CUG-1840"  
 /codon\_start 1  
 /translation\_except\_pos (1771..1771,aa:Met)  
 /product "nuclear hormone receptor PRK2-C"  
 /protein\_id "CA55493.1"  
 /db\_xref "GI:5852068"  
 /translation "MDPRGEVCAKMLPNRSPRGPEANLHYPRDSNMTDFVCEITE  
 SVKQSVNADFEVGGPQICRVGSGKATGHPNWTCEGKCFRRAMKRNALRCP  
 KQACETFRTERQACALRKLKESKSNEMINSDEVERRALIRKRSKRTGTP  
 LQVGLTFEPMQMTRELMDQMTPTTFESHKPNVSIQJGCDSDSVNKKPAAG  
 QKELFSLPHMAMSTYMKGISPAKVISYRDLPIEDQISLKGAEFLQGLRNT  
 VEMALNEMLSGLSYGLIDIAQGFQDILIEPMKLFYMLKQIDHEEVLMQALST  
 ESRPGVLDHREVLQDFQALIKSYLENRPAPHRFLFLKIMMTELSINAST  
 HIGRLRIQDHPAPRPLMGELFGITGS"  
 1840..4013  
 /gene "PRK2"  
 /function "transcription factor"  
 /note "full length receptor"  
 /codon\_start 1  
 /product "nuclear hormone receptor PRK2-A"  
 /protein\_id "CA55492.1"  
 /db\_xref "GI:5852067"  
 /translation "MEVRKPSMNHTPVECEDESQVGRKSVNADFEVGGPQICVC  
 GKATGHPNWTCEGKCFRRAMKRNALRCPFRKACETFRTERQACALRKLK  
 LESKSNEMINSDEVERRALIRKRSKRTGTPQVGLTFEPMQMTRELMDQMT  
 PTTFESHKPNVSIQJGCDSDSVNKKPAAGQKELFSLPHMAMSTYMKGIS  
 SPAKVISYRDLPIEDQISLKGAEFLQGLRNTVEMALNEMLSGLSYGLIDIA  
 QGFQDILIEPMKLFYMLKQIDHEEVLMQALSTESRPGVLDHREVLQDFQAL  
 IKSYLENRPAPHRFLFLKIMMTELSINASTHIGRLRIQDHPAPRPLMGELF  
 GITGS"  
 2005..4013  
 /gene "PRK2"  
 /note "truncated receptor"  
 /codon\_start 1  
 /product "nuclear hormone receptor PRK2-B"  
 /protein\_id "CA55494.1"  
 /db\_xref "GI:5852069"  
 /translation "MTRECKGPPRRAMKRNALRCPKFACEITKTRSGQACRL  
 KCTLESKSNEMINSDEVERRALIRKRSKRTGTPQVGLTFEPMQMTRELMDQ  
 MTPTTFESHKPNVSIQJGCDSDSVNKKPAAGQKELFSLPHMAMSTYMKGIS  
 SPAKVISYRDLPIEDQISLKGAEFLQGLRNTVEMALNEMLSGLSYGLIDIA  
 QGFQDILIEPMKLFYMLKQIDHEEVLMQALSTESRPGVLDHREVLQDFQAL  
 IKSYLENRPAPHRFLFLKIMMTELSINASTHIGRLRIQDHPAPRPLMGELF  
 GITGS"  
 BASE COUNT 1134 a 1105 c 1056 g 1042 t  
 ORIGIN  
 Query Match 80.4% Score 1846.6 DB 9 Length 4337  
 Best local similarity 94.6% Pred. No. 0  
 Matches 2011 Conservative 0 Mismatches 4 Indels 112 Gaps 27  
 UY 1 TAAATATAGTGAAGAAAGATATCTATCTGGGAAATATCACTATCACTAG 60  
 DB 1 TAAATAATAGTGAAGAAAGATATCTATCTATCTATCTATCTATCTATCT 1596  
 UY 61 GATCGAAGAAAGAAAGATCTCTTAACTATGAGTGGAAATCTCGACCTGCAAGCC 120  
 DB 1597 GATCGAAGAAAGAAAGATCTCTTAACTATGAGTGGAAATCTCGACCTGCAAGCC 1656  
 UY 121 AAGTATTCAGATGAAGAAAGAAAGATATAGTAAATGCTGCTGAGCAAGGCGAG 180  
 DB 1657 AAGTATTCAGATGAAGAAAGAAAGATATAGTAAATGCTGCTGAGCAAGGCGAG 1716  
 UY 181 CTGAAGCTCTGATGAAGCTACTCTTGAATCTGATCTGATCTGATCTGATCT 240

1717 GAGGCTCTTGGTAAAGTACTCTTGAATCAATCTCTATATCAATCTGTTAAAGTCAAT 1776  
 UY 241 CCGAGGAAAGAAAGTCAAGGCTCAAGAAATATCACTCAAGGCTCAAGGCTCAAGGCTCA 600  
 DB 1777 CCGAGGAAAGAAAGTCAAGGCTCAAGAAATATCACTCAAGGCTCAAGGCTCAAGGCTCA 1846  
 UY 401 AAGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 640  
 DB 1847 AAGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 1896  
 UY 361 AAGAAATCTATCTGAGAAAGGCTATCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 420  
 DB 1897 ACACAGCTCTGCTTCTTGAAGAAAGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 1956  
 UY 421 ATCTCTCTATATCTGAGGAAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 480  
 DB 1957 ATCTCTCTATATCTGAGGAAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 2016  
 UY 481 GAGTCAAGGCTCTTCTGAGGAAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 540  
 DB 2017 GATATCAAGGCTCTTCTGAGGAAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 2076  
 UY 541 CAGGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 600  
 DB 2077 CCGAGGAAAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 2136  
 UY 601 CCGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 660  
 DB 2137 CCGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2196  
 UY 661 GAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAG 720  
 DB 2197 GAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAG 2256  
 UY 721 GTCGAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 780  
 DB 2257 GTCGAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGG 2316  
 UY 781 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 840  
 DB 2317 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2376  
 UY 841 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 900  
 DB 2358 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2416  
 UY 901 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 960  
 DB 2358 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2416  
 UY 961 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 1020  
 DB 2386 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2445  
 UY 1021 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 1080  
 DB 2446 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2505  
 UY 1081 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 1140  
 DB 2506 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2565  
 UY 1141 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 1200  
 DB 2566 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2625  
 UY 1201 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 1260  
 DB 2626 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 2685  
 UY 1261 AAGGCTCTGAGGCTCTGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 1320



[illegible]





[illegible][illegible]

[illegible][illegible]



| Matches | 12/5 | Conservative  | 0    | Mismatches | 292 | Indels | 5 | Gaps |
|---------|------|---|------|------------|-----|--------|---|------|
| 07      | 203  | CTTATTCGATCTCTTGATCGGATTTCTTCAAGTGTGATCCCTCGGAGGAATTCAGACAA       | 262  |            |     |        |   |      |
| 08      | 44   | CTGGAGCAACAACCATCTCTCTCAATCTTCAAGCTGCACCTGCGGAAATCCCAACAAA        | 103  |            |     |        |   |      |
| 09      | 263  | GAAGCTTACCACTCAAGTGTCTCAAGGCGCCGAAAGCAAGCTTCAAGTGTAGAGCCCAAG      | 352  |            |     |        |   |      |
| 10      | 104  | AGCACTTATCGCCGACAGACTCTGAGCACTACAGATGTAACCTCTAGACATGAGACTGAGG     | 153  |            |     |        |   |      |
| 07      | 423  | AAAGCTTGAAGCTGATGATCTTTGATCTGTGTGAAGATCGGAATCTGTCTCTCTGGAAGC      | 382  |            |     |        |   |      |
| 08      | 164  | AGCACTTACACCCACATCTTCCCTCTTGACATCTGAAAGACAGACTCTGCTTGGAAACAC      | 223  |            |     |        |   |      |
| 09      | 383  | CTGATCTGTCAAGCGAATTTAGGAAGTCAAGATCTCCCAATCTCTCGATATCTAGAGAAC      | 442  |            |     |        |   |      |
| 10      | 224  | CTCA---CTCAACGTAGAGAGAGAGAGATGAGAGCTCTCAATCTTCCCGTGTATGTCGGACA    | 280  |            |     |        |   |      |
| 09      | 443  | GAAGCTCTGTCT      | 502  |            |     |        |   |      |
| 10      | 281  | AGCCCTCATTCCTTACCATCTCAATGTATGATACCTGTGAAGACATGTAAGAGCTTTTTCGAA   | 340  |            |     |        |   |      |
| 09      | 503  | AGGCTATGAAGCGAAGCG        | 562  |            |     |        |   |      |
| 10      | 341  | AGGCTATGAAGCG       | 400  |            |     |        |   |      |
| 09      | 563  | CGCGGATGATCG        | 622  |            |     |        |   |      |
| 10      | 401  | CGCGGATGATCG        | 460  |            |     |        |   |      |
| 09      | 623  | CTGAAGAGAGATGATGATCT      | 682  |            |     |        |   |      |
| 10      | 461  | CTGAAGAGAGATGATGATCT      | 520  |            |     |        |   |      |
| 09      | 683  | GAAGAGAGATGATGATCT        | 742  |            |     |        |   |      |
| 10      | 521  | GAAGAGAGATGATGATCT        | 580  |            |     |        |   |      |
| 09      | 743  | AGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 802  |            |     |        |   |      |
| 10      | 581  | AGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 640  |            |     |        |   |      |
| 09      | 803  | CGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT      | 862  |            |     |        |   |      |
| 10      | 641  | CGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT      | 700  |            |     |        |   |      |
| 09      | 863  | CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 922  |            |     |        |   |      |
| 10      | 701  | CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 760  |            |     |        |   |      |
| 09      | 923  | GTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT   | 982  |            |     |        |   |      |
| 10      | 761  | GTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT   | 820  |            |     |        |   |      |
| 09      | 983  | CGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 1042 |            |     |        |   |      |
| 10      | 821  | CGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 880  |            |     |        |   |      |
| 09      | 1103 | GAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 1102 |            |     |        |   |      |
| 10      | 881  | GAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 940  |            |     |        |   |      |
| 09      | 1103 | CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 1162 |            |     |        |   |      |
| 10      | 941  | CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT     | 1000 |            |     |        |   |      |
| 09      | 1163 | GAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 1222 |            |     |        |   |      |
| 10      | 1001 | GAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 1060 |            |     |        |   |      |
| 09      | 1223 | GTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT   | 1282 |            |     |        |   |      |
| 10      | 1061 | GTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT   | 1120 |            |     |        |   |      |

[illegible]

```

RESULT 14
AF151377 LOCUS 1755 bp mRNA ROD 05-AUG-1999
DEFINITION Rattus norvegicus pregnane X receptor mRNA, complete cds.
ACCESSION AF151377
VERSION AF151377.1 GI:5702212
KEYWORDS
SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 1755)
Zheng,H., Lechlyse,E., Liu,L., Hu,M., Maloney,L., Zhu,W., and Yan,B.
Rat Pregnane X Receptor: Molecular Cloning, Tissue Distribution,
and Xenobiotic Regulation
Archive of Biochemistry and Biophysics 368 (1), 14-22 (1999)
1415106
2 (bases 1 to 1755)
Zheng,H., Lechlyse,E., Liu,L., Hu,M., Maloney,L., Zhu,W., and Yan,B.
Direct Submission
Submitted (13-MAY-1999) Biomedical Sciences, University of Rhode
Island, 41 Lower College Road, Kingston, RI 02881, USA
Location/Comments
1..1755
FEATURES
SOURCE
/organism "Rattus norvegicus"
/strain "Sprague-Dawley"
/db_xref "taxon:10116"
218..1513
/region_start 1
/product "pregnane X receptor"
/protein_id "AAD4214.1"
/db_xref "GI:5702212"
/translation "MRPEFNNHNYGLVGRFALSVLEFPINVEEDKRLQTFVWQIDK
ANVHYNMTCCCKCPFKKAKMKENYVLRPEKCTETFKIKIQVQACR;KGLTES
GKKKIKNSIAVAGQRRLIKKKRKKLEKRPNGTGLFEQNALQPLMAQVQLEP
TLESHKRPGLPAVHSDGLPYVALSDLELALATSSQIMKSNVYMKISVQKEDDPS
LNNYQPSKSKCKEKLPIPLDVAQNSVYKRVKAKTSHKELILEQISLKCA
TEEMLEKRNPTETETWVGGLAVFEDNGGQGLLPLMKRPLKRLQLEP

```





Sat Mar 2 12:27:38 2002

us-09-276-935b-13.rge

Page

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: March 2, 2002, 00:48:49 ; Search time 139.68 Seconds

Title: US-09-276-4358-13  
Port of origin: 2146

Sequence: 1 tgaatataatg ttaagagca.....aaatgcatagtatctgtac 2146

Scoring table: INFINITY\_NUC

Searched: 930621 soys, 428662619 residues

Total number of hits satisfy any chosen parameters: 1861242

Maximum DB seq length: 20000000000

Post-processing: Minimum Match (9)

### Listing first 45 summaries

|     |  |
|-----|--|
| 1:  | /S10S2/qcudata/geneseq/geneseqn/NA1960.DAT.* |
| 2:  | /S10S2/qcudata/geneseq/geneseqn/NA1961.DAT.* |
| 3:  | /S10S2/qcudata/geneseq/geneseqn/NA1962.DAT.* |
| 4:  | /S10S2/qcudata/geneseq/geneseqn/NA1963.DAT.* |
| 5:  | /S10S2/qcudata/geneseq/geneseqn/NA1964.DAT.* |
| 6:  | /S10S2/qcudata/geneseq/geneseqn/NA1965.DAT.* |
| 7:  | /S10S2/qcudata/geneseq/geneseqn/NA1966.DAT.* |
| 8:  | /S10S2/qcudata/geneseq/geneseqn/NA1967.DAT.* |
| 9:  | /S10S2/qcudata/geneseq/geneseqn/NA1968.DAT.* |
| 10: | /S10S2/qcudata/geneseq/geneseqn/NA1969.DAT.* |
| 11: | /S10S2/qcudata/geneseq/geneseqn/NA1970.DAT.* |
| 12: | /S10S2/qcudata/geneseq/geneseqn/NA1971.DAT.* |
| 13: | /S10S2/qcudata/geneseq/geneseqn/NA1972.DAT.* |
| 14: | /S10S2/qcudata/geneseq/geneseqn/NA1973.DAT.* |
| 15: | /S10S2/qcudata/geneseq/geneseqn/NA1974.DAT.* |
| 16: | /S10S2/qcudata/geneseq/geneseqn/NA1975.DAT.* |
| 17: | /S10S2/qcudata/geneseq/geneseqn/NA1976.DAT.* |
| 18: | /S10S2/qcudata/geneseq/geneseqn/NA1977.DAT.* |
| 19: | /S10S2/qcudata/geneseq/geneseqn/NA1978.DAT.* |
| 20: | /S10S2/qcudata/geneseq/geneseqn/NA1979.DAT.* |
| 21: | /S10S2/qcudata/geneseq/geneseqn/NA2000.DAT.* |
| 22: | /S10S2/qcudata/geneseq/geneseqn/NA2001.DAT.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARY

| Result No. | Score  | Query Match | Length | OR | ID        | Description        |
|------------|--------|-------------|--------|----|-----------|--------------------|
| 1          | 2146   | 100.0       | 2146   | 20 | AAZ079497 | Human pregnane X r |
| 2          | 2115.4 | 98.6        | 5057   | 20 | AAZ594975 | Ena encoding an in |
| 3          | 2101.4 | 97.9        | 4093   | 20 | AAZ788070 | Human nkr7 cDNA.   |
| 4          | 2075.2 | 96.7        | 2910   | 20 | AAZ56242  | Human vitamin D r  |
| 5          | 2006.4 | 94.5        | 4243   | 20 | AAZ59374  | Ena encoding an in |
| 6          | 1866.6 | 87.0        | 2802   | 20 | AAZ56244  | Human vitamin D r  |
| 7          | 1856.6 | 87.0        | 2850   | 20 | AAZ78808  | Human nkr7-1 cDNA. |
| 8          | 1744.4 | 81.3        | 2068   | 20 | AAZ89680  | Human steroid and  |
| 9          | 1744.4 | 81.3        | 2068   | 22 | AAH25487  | Nucleotide sequen  |
| 10         | 1474   | 64.0        | 1474   | 20 | AAZ59467  | Ena encoding an in |
| 11         | 1430.2 | 62.0        | 1422   | 20 | AAZ59468  | DNA encoding an in |

[illegible]



[illegible][illegible]



|    |      |  |      |
|----|------|--|------|
| Db | 472  | aaatctgaagatctcccaaatctacacagatataatgaggaagaagccactgcatctctca    | 533  |
| Gy | 464  | atctctatgaatgagtgaaatgcatgcaatgacctctcttcagagagagcatgaatacgaaccc | 523  |
| Db | 532  | atctctatgaatgagtgaaatgcatgcaatgacctctcttcagagagagcatgaatacgaaccc | 591  |
| Gy | 524  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 583  |
| Db | 592  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 651  |
| Gy | 584  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 643  |
| Db | 652  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 711  |
| Gy | 644  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 703  |
| Db | 712  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 771  |
| Gy | 704  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 763  |
| Db | 772  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 831  |
| Gy | 764  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 823  |
| Db | 832  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 891  |
| Gy | 824  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 883  |
| Db | 892  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 951  |
| Gy | 884  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 943  |
| Db | 952  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1011 |
| Gy | 944  | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1003 |
| Db | 1012 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1071 |
| Gy | 1004 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1063 |
| Db | 1072 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1131 |
| Gy | 1064 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1123 |
| Db | 1132 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1191 |
| Gy | 1124 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1183 |
| Db | 1192 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1251 |
| Gy | 1184 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1243 |
| Db | 1252 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1311 |
| Gy | 1244 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1303 |
| Db | 1312 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1371 |
| Gy | 1304 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1363 |
| Db | 1372 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1431 |
| Gy | 1364 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1423 |
| Db | 1432 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1491 |
| Gy | 1424 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1483 |
| Db | 1492 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1551 |
| Gy | 1484 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1543 |
| Db | 1552 | agctcaaatgagacctctccgaagaagagcctgagatcaaccacgaagaagccagacagat    | 1611 |

[illegible]





[illegible][illegible]







[illegible][illegible]

XX Sequence 2068 BP: 520 A: 541 G: 586 Q: 420 T: 1 Other:

Query Match: 81.4% Score 1744.4; DB 22: Length 2068;  
 Best Local Similarity: 99.48; Pred. No. 0;  
 Matches 1761: Conservative 0; Mismatches 7; Indels 3; Gaps 3;

UY 1 tgaatataatgaagaagaagaatgtctcatatccggaagaatcataactatgaactag 60  
 Db 1aaadattagtgagagagaaadattgtctcatatccggaagaatcataactatgaactag 339  
 UY 61 gacag 120  
 Db 340 gacag 399  
 UY 121 aagatgttctcaatgtatataaagaagaagaataagtaactcctgtctctcaagaagag 180  
 Db 400 aagatgttctcaatgtatataaagaagaagaataagtaactcctgtctctcaagaagag 459  
 UY 181 cgaatccttgaataaactcctctctctctctctctctctctctctctctctctctct 240  
 Db 460 cgaatccttgaataaactcctctctctctctctctctctctctctctctctctctct 519  
 UY 241 cctcaag 300  
 Db 520 cctcaag 579  
 UY 301 aacctgaag 360  
 Db 580 aacctgaag 639  
 UY 361 acgaatctgtctctgaagaagagagagagagagagagagagagagagagagagag 420  
 Db 640 acgaatctgtctctgaagaagagagagagagagagagagagagagagagagagag 699  
 UY 421 atctctcag 480  
 Db 700 atctctcag 759  
 UY 481 ggaatctgaagaag 540  
 Db 760 ggaatctgaagaag 819  
 UY 541 cgaagaag 600  
 Db 820 cgaagaag 879  
 UY 601 cgaagaatctcctgaagaagagagagagagagagagagagagagagagagagag 660  
 Db 880 cgaagaatctcctgaagaagagagagagagagagagagagagagagagagagag 919  
 UY 661 gaaag 720  
 Db 940 gaaag 999  
 UY 721 gtaag 780  
 Db 1000 gtaag 1059  
 UY 781 aaaaacttgaagaatcctctcctctctctctctctctctctctctctctctctct 840  
 Db 1060 aaaaacttgaagaatcctctcctctctctctctctctctctctctctctctctct 1119  
 UY 841 aatgtctgtcaatgttcaagaatcctctctctctctctctctctctctctctctct 900  
 Db 1120 aatgtctgtcaatgttcaagaatcctctctctctctctctctctctctctctctct 1179  
 UY 901 agcgaag 959  
 Db 1180 agcgaag 1239  
 UY 960 tggcagatgtctctgaagaatcacaaccccccaagaagaatgtggagagagagag 1019

Db 1240 tggcagatgtctctgaagaatcacaaccccccaagaagaatgtggagagagagag 1298  
 UY 1020 gctggcgaag 1079  
 Db 1299 gctggcgaag 1358  
 UY 1080 aatgaatcctcctactcgaagaatctcctcctcctcctcctcctcctcctcctcct 1139  
 Db 1359 aatgaatcctcctactcgaagaatctcctcctcctcctcctcctcctcctcctcct 1418  
 UY 1140 cgaatctcgaag 1199  
 Db 1419 cgaatctcgaag 1478  
 UY 1200 ggaatgtgag 1259  
 Db 1479 ggaatgtgag 1538  
 UY 1260 actgaagaag 1319  
 Db 1539 actgaagaag 1598  
 UY 1320 gtaatgtatgaagaag 1379  
 Db 1599 gtaatgtatgaagaag 1658  
 UY 1380 cgaagaatgtgaagaag 1439  
 Db 1659 cgaagaatgtgaagaag 1718  
 UY 1440 caatcgaag 1499  
 Db 1719 caatcgaag 1778  
 UY 1500 gctcgaag 1559  
 Db 1779 gctcgaag 1838  
 UY 1560 ctttctatgag 1619  
 Db 1839 ctttctatgag 1898  
 UY 1620 tgaatgaag 1679  
 Db 1898 tgaatgaag 1958  
 UY 1680 aagaagaatgtgaagaatcctcctcctcctcctcctcctcctcctcctcctcct 1739  
 Db 1958 aagaagaatgtgaagaatcctcctcctcctcctcctcctcctcctcctcctcct 2017  
 UY 1740 tctctctatgaagaatcctcctcctcctcctcctcctcctcctcctcctcctcct 1799  
 Db 2018 tctctctatgaagaatcctcctcctcctcctcctcctcctcctcctcctcctcct 2068

RESULT 10  
 AAX59967  
 ID AAX59967 standard; DNA: 1474 BP.  
 AC AAX59967;  
 XX  
 DT 04-AUG-1999 (first entry)  
 XX  
 DE DNA encoding an intranuclear receptor protein.  
 KW Human: intranuclear receptor protein; drug development; diagnosis;  
 KW treatment; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN JP11127872-A.  
 XX











Core version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

DM nucleic acid search, using SW model

Run on: March 2, 2002, 00:47:59 ; Search time 55.15 seconds

Title: HS-09-276-935H-13

Sequence: 1 TGAAGATATAGGTTGAGGAGAG...AGGTGCATGATATCTCTG 2146

Scoring table: IDENTITY\_NDC

Current [0] = (Anext - 1)

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Maximum PB seq length: 200000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

```
Database : Issued/Issued_NA:*
1 : /cgn2.6/prodata/2/ina/5A.COMM.seq.*
2 : /cgn2.6/prodata/2/ina/5B.COMM.seq.*
3 : /cgn2.6/prodata/2/ina/6A.COMM.seq.*
4 : /cgn2.6/prodata/2/ina/6B.COMM.seq.*
5 : /cgn2.6/prodata/2/ina/PCrTOS.COMM.seq.*
6 : /cgn2.6/prodata/2/ina/backfilltest.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Quarry | Match | Length | DB                | LD                 | Description |
|------------|-------|--------|-------|--------|-------------------|--------------------|-------------|
| 1          | 746.6 | 34.3   | 801   | 4      | US-05-276-531-111 | Sequence 11, Appl  |             |
| 2          | 278.6 | 13.0   | 2043  | 1      | US-07-717-7361-6  | Sequence 6, Appl   |             |
| 3          | 272.8 | 12.7   | 1399  | 1      | US-07-737-7361-5  | Sequence 5, Appl   |             |
| 4          | 170.6 | 7.9    | 1450  | 1      | US-08-459-489-1   | Sequence 1, Appl   |             |
| 5          | 170.6 | 7.9    | 1450  | 1      | US-08-458-686-1   | Sequence 1, Appl   |             |
| 6          | 170.6 | 7.9    | 1450  | 1      | US-07-843-550C-1  | Sequence 1, Appl   |             |
| 7          | 170.6 | 7.9    | 1450  | 5      | PCI-US94-01559-1  | Sequence 1, Appl   |             |
| 8          | 97.6  | 4.5    | 1898  | 1      | US-08-342-6119-1  | Sequence 1, Appl   |             |
| 9          | 97.6  | 4.5    | 1979  | 2      | US-08-669-619B-1  | Sequence 1, Appl   |             |
| 10         | 97.6  | 4.5    | 2030  | 1      | US-08-340-518-1   | Sequence 1, Appl   |             |
| 11         | 97.6  | 4.5    | 2030  | 1      | US-08-340-283-1   | Sequence 1, Appl   |             |
| 12         | 97.6  | 4.5    | 2030  | 2      | US-08-646-248-1   | Sequence 1, Appl   |             |
| 13         | 97.6  | 4.5    | 2030  | 5      | PCI-US95-11924-1  | Sequence 1, Appl   |             |
| 14         | 97.6  | 4.5    | 2030  | 5      | PCI-US95-13931-1  | Sequence 1, Appl   |             |
| 15         | 97.2  | 4.5    | 1813  | 5      | PCI-US94-12883-3  | Sequence 3, Appl   |             |
| 16         | 91.4  | 4.3    | 2928  | 2      | US-08-095-7288-3  | Sequence 3, Appl   |             |
| 17         | 91.4  | 4.3    | 2928  | 5      | PCI-US92-02320A-3 | Sequence 3, Appl   |             |
| 18         | 91.4  | 4.3    | 2040  | 6      | 5176171-1         | Patent No. 5176177 |             |
| 19         | 91.4  | 4.3    | 3036  | 1      | US-08-306-091A-52 | Sequence 52, Appl  |             |
| 20         | 91.4  | 4.3    | 3036  | 2      | US-08-095-7281-1  | Sequence 1, Appl   |             |
| 21         | 91.4  | 4.3    | 4036  | 5      | PCI-US92-02320A-1 | Sequence 1, Appl   |             |
| 22         | 91.4  | 4.3    | 3511  | 3      | US-08-842-747-13  | Sequence 13, Appl  |             |
| 23         | 89.8  | 4.2    | 704   | 2      | US-08-592-383-5   | Sequence 5, Appl   |             |
| 24         | 89.8  | 4.2    | 2040  | 2      | US-08-592-383-1   | Sequence 1, Appl   |             |
| 25         | 85.2  | 4.0    | 1934  | 4      | US-08-776-844-1   | Sequence 1, Appl   |             |
| 26         | 85.2  | 4.0    | 1959  | 1      | US-08-342-611A-3  | Sequence 3, Appl   |             |
| 27         | 85.2  | 4.0    | 1959  | 5      | PCI-US94-12883-4  | Sequence 4, Appl   |             |

|    |      |     |       |   |                      |                     |
|----|------|-----|-------|---|----------------------|---------------------|
| 26 | 83.8 | 4.9 | 2.652 | 2 | US-08-59.2-48.3-2    | Sequence 4, Aff1    |
| 29 | 80.8 | 4.8 | 1.860 | 2 | US-08-47.2-65.2-7    | Sequence 7, Aff1    |
| 30 | 80.8 | 4.8 | 1.860 | 5 | PcT1-US95-16.411-7   | Sequence 7, Aff1    |
| 31 | 79.6 | 4.7 | 1.659 | 1 | US-08-33.3-35.8-7    | Sequence 7, Aff1    |
| 32 | 79.6 | 4.7 | 1.659 | 1 | US-08-44.3-69.4-7    | Sequence 7, Aff1    |
| 33 | 79.6 | 4.7 | 1.659 | 1 | US-08-65.4-50.1-7    | Sequence 7, Aff1    |
| 34 | 77.6 | 4.6 | 2.970 | 5 | PcT1-US92.2 (66.5)-1 | Sequence 1, Aff1    |
| 35 | 77.7 | 4.6 | 1.688 | 2 | US-08-64.9-61.0-8-2  | Sequence 1, Aff1    |
| 36 | 74.8 | 4.5 | 1.649 | 2 | US-08-46.6-120-1     | Sequence 1, Aff1    |
| 37 | 74.8 | 4.5 | 1.649 | 5 | PcT1-US94-07.2-66-3  | Sequence 1, Aff1    |
| 38 | 74.4 | 4.5 | 1.576 | 6 | 5.6504.3-2-1         | Factor N1, 5.6504.3 |
| 39 | 74.4 | 4.5 | 2.241 | 3 | US-09-14.4-75.9-17   | Sequence 17, Aff1   |
| 40 | 74.4 | 4.5 | 2.295 | 3 | US-09-14.4-75.9-19   | Sequence 19, Aff1   |
| 41 | 74.4 | 4.5 | 2.303 | 3 | US-09-14.4-75.9-21   | Sequence 21, Aff1   |
| 42 | 73.2 | 4.4 | 2.686 | 1 | US-08-33.3-45.8-11   | Sequence 11, Aff1   |
| 43 | 73.2 | 4.4 | 2.686 | 1 | US-08-46.3-69.4-11   | Sequence 11, Aff1   |
| 44 | 73.2 | 4.4 | 2.686 | 1 | US-08-59.4-50.1-11   | Sequence 11, Aff1   |
| 45 | 72.8 | 4.4 | 2.687 | 4 | US-07-10.4-194-11    | Sequence 11, Aff1   |

## AL.1 (NMHS)

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



[illegible]

RES0311 3  
US-07-747-368-5  
Sequence 5, Application US/07747368  
Patent No. 5260199  
GENERAL INFORMATION:  
APPLICANT: Deluca, Hector F.  
APPLICANT: Ross, Troy R.  
APPLICANT: Prahl, Jean M.  
TITLE OF INVENTION: Method of Producing  
NUMBER OF INVENTION: 1,25-Dihydroxyvitamin D3 Receptor Protein  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carl R. Schwartz, Esq., c/o Quarles & Brady  
STREET: 411 East Wisconsin Avenue

[illegible]











```

Sequence 1, Application US/08440283
Patent No. 5679518
GENERAL INFORMATION:
Applicant: Fritzelman, Eitan
Applicant: Hollaway, M. Katharine
Applicant: Rodan, Gideon
Applicant: Schmidt, Suzi Jane
Applicant: Solomon, Azriel
Applicant: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08440283
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DCKET NUMBER: 19327
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-310-283-1

```

```

Query Match 4.5% Score 97.6; DB 1; Length 2030;
Best Local Similarity 54.8%; Pred. No. 2.6e-14;
Matches 215; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

```

```

QY 1 ctgaatttgaacatgaagacacgaagatctatctctgaagaccccaagtcgaacgca 397
DB 417 CACACTGGCTGATCCATCCGGAAGAACACGAGACCGGCAAGCAAGAAAGAGCCGCA 476
QY 498 atgaagatgaagatgaagatgaagatgaagatgaagatgaagatgaagatgaag 457
DB 477 CCGGGAAGATGCTGGGACACGAGCTTTAGCGTGTGTGGGAAACAGGCTTCGGCTTC 556
QY 458 actcaatgtcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 517
DB 537 ACTACAACTGCTCACTGACGTCGGAAGCTTCGCAAGCGCTCTCCGCGCAGAGTGGT 596
QY 518 aagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 574
DB 597 GTGGGGGTAAGCTGATGCTGCGGGGGTGGGGGGAACCTGCGAGATGAGACCTTTAT 656
QY 575 ggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 634
DB 657 GCGCGAAGTGGCGAGCAGTGTGGGCTGGCAAGTGCAGAGGCGAGGATGAGGAGAGT 716
QY 645 tgaatgtcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 604
DB 717 GCGTCTTTCTGGAAGAACAGATTCGGAAGAGAGATTCGGAACACGAGAGAGAGT 776

```

```

QY 695 aagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 726
DB 777 CACACTGGCTGATCCATCCGGAAGAACACGAGACCGGCAAGCAAGAAAGAGCCG 808

```

```

RESULT 12
US-08-646-248-1
Sequence 1, Application US/08646248
Patent No. 5939322
GENERAL INFORMATION:
Applicant: Fritzelman, Eitan
Applicant: Hollaway, M. Katharine
Applicant: Rodan, Gideon
Applicant: Schmidt, Suzi Jane
Applicant: Solomon, Azriel
Applicant: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,248
FILING DATE: 14-MAY-1996
CLASSIFICATION: 516
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/330,284
FILING DATE: 27-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DCKET NUMBER: 19327
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-646-248-1

```

```

Query Match 4.5% Score 97.6; DB 2; Length 2030;
Best Local Similarity 54.8%; Pred. No. 2.6e-14;
Matches 215; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

```

```

QY 338 ctgaatttgaacatgaagacacgaagatctatctctgaagaccccaagtcgaac 397
DB 417 CACACTGGCTGATCCATCCGGAAGAACACGAGACCGGCAAGCAAGAAAGAGCCGCA 476
QY 398 atgaagatgaagatgaagatgaagatgaagatgaagatgaagatgaagatgaag 457
DB 477 CCGGGAAGATGCTGGGACACGAGCTTTAGCGTGTGTGGGAAACAGGCTTCGGCTTC 556
QY 458 actcaatgtcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 517
DB 537 ACTACAACTGCTCACTGACGTCGGAAGCTTCGCAAGCGCTCTCCGCGCAGAGTGGT 596

```



```

CY 348 ctgaattatagatgagagagagatctatctcctggaagcccaatgacacgag 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 417 CAGACGCGGCTGATGCGGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 398 atgagagagatcagagatcagagatcagagatcagagatcagagatcagag 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 477 CCGGAGAGATGCTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 458 acctcaatgagatgagatgagatgagatgagatgagatgagatgagatgag 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 517 ACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 518 agcagagagatgagatgagatgagatgagatgagatgagatgagatgagat 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 597 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 575 gacagagagatgagatgagatgagatgagatgagatgagatgagatgagat 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 657 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 635 tgaatgagatgagatgagatgagatgagatgagatgagatgagatgagat 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 717 GGTCTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 717 gacagagagatgagatgagatgagatgagatgagatgagatgagatgagat 776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 777 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 15

PCT-US94-12883-3

Sequence 3, Application PC/US9412883

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: OBLIGATIONS NUCLEAR RECEPTOR: COMPOSITIONS AND

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSER: Arnold White &amp; Burke

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: United States of America

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

SOFTWARE: Patent in Release #1.0, Version

SOFTWARE: #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/12883

FILING DATE: Concurrently herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/152,003

FILING DATE: 10-Nov-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: BARBARA S. KIRKELL

REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: ARCD154P--

TELEPHONE: (512) 418-1000

TELEFAX: (713) 789-2679

TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1813 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US94-12883-3

Query Match

Best local similarity: 58.8%, Score 97.2, DB 5, Length 1813

Matches 187, Conservative 0, Mismatches 128, Indels 3, Gaps 1

```

CY 412 gctcccaaatctgagatgagatgagatgagatgagatgagatgagatgagat 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 229 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 472 acatgagagatgagatgagatgagatgagatgagatgagatgagatgagat 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 AGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 532 tgcctcttgcagagagagagagagagagagagagagagagagagagagagag 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 TATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 589 gctcagagatgagatgagatgagatgagatgagatgagatgagatgagat 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 409 CAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 649 gacagagagatgagatgagatgagatgagatgagatgagatgagatgagat 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 469 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CY 709 gacagagatgagatgagatgagatgagatgagatgagatgagatgagat 726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 529 CAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: March 2, 2002, 01:50:22

Job time: 3743 sec

1

...













[illegible]

|                       |      |  |                                      |            |                         |             |   |      |
|-----------------------|------|--|--------------------------------------|------------|-------------------------|-------------|---|------|
| Gy                    | 2045 | ctatggttgcctcctccacagctggcttcctccctctctcctgactgattatgaa  |                                      |            |                         |             |   | 2104 |
| Dd                    | 121  | tttatatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt   |                                      |            |                         |             |   | 180  |
| Gy                    | 2105 | ctccagctctgacctaccgaatacaatgatcatctctatcg  |                                      |            |                         |             |   | 2146 |
| Dd                    | 181  | ctccacaccccttgacctatgccatctctatcacatctctctg  |                                      |            |                         |             |   | 222  |
| RESULT                | 9    | N94132   | 454 bp                               | mRNA       | EST                     | v5-Apr-1996 |   |      |
| LOCUS                 |      | N94132   | ZAG500..J1 Soares fetal liver spleen | INTLS      | Homo sapiens cDNA clone |             |   |      |
| DEFINITION            |      | IMAGE:ZAG500.57 mRNA sequence.   |                                      |            |                         |             |   |      |
| ACCESSION             |      | N94132   | GI:1266441                           |            |                         |             |   |      |
| VERSION               |      | N94132.1   | GI:1266441                           |            |                         |             |   |      |
| KEYWORDS              |      | EST  |                                      |            |                         |             |   |      |
| SOURCE                |      | human  |                                      |            |                         |             |   |      |
| ORGANISM              |      | Homo sapiens   |                                      |            |                         |             |   |      |
| REFERENCE             |      | Bokoyeva M; Mezoez C; Chordata; Craniata; Vertebrata; Euteleostomi;                                      |                                      |            |                         |             |   |      |
| AUTHORS               |      | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.   |                                      |            |                         |             |   |      |
|                       |      | 1 (bases 1 to 454)   |                                      |            |                         |             |   |      |
|                       |      | Hillier L., Clark N., Duboucq T., Elliston K., Hawks M., Holman  |                                      |            |                         |             |   |      |
|                       |      | M., Holman R., Kudva I., Le M., Lemon G., Merida B., Parsons J.,   |                                      |            |                         |             |   |      |
|                       |      | Ritkin D., Rothberg J., Soares M., Tan P., Trevaskis E., Watson  |                                      |            |                         |             |   |      |
|                       |      | K., Williamson A., Wohldmann P. and Wilson R.  |                                      |            |                         |             |   |      |
|                       |      | The WABSI-Merck EST Project  |                                      |            |                         |             |   |      |
|                       |      | Unpublished (1995)   |                                      |            |                         |             |   |      |
|                       |      | Contact : Wilson RK  |                                      |            |                         |             |   |      |
|                       |      | Washington University School of Medicine   |                                      |            |                         |             |   |      |
|                       |      | 4444 Forest Park Parkway, Box 8091, St. Louis, MO 63108  |                                      |            |                         |             |   |      |
|                       |      | Telex: 314 286 1800  |                                      |            |                         |             |   |      |
|                       |      | Fax: 314 286 1810  |                                      |            |                         |             |   |      |
|                       |      | Email: est.watson@wustl.edu  |                                      |            |                         |             |   |      |
|                       |      | This clone is available royalty-free through LNCx. Contact the   |                                      |            |                         |             |   |      |
|                       |      | IMAGE Consortium ( <a href="mailto:info@image.lnc.gov">info@image.lnc.gov</a> ) for further information. |                                      |            |                         |             |   |      |
|                       |      | Seq primer: mob.MFGA-E1  |                                      |            |                         |             |   |      |
|                       |      | High quality sequence stop: 416.   |                                      |            |                         |             |   |      |
| FEATURES              |      | Location/Qualifiers  |                                      |            |                         |             |   |      |
| SOURCE                |      | 1 454  |                                      |            |                         |             |   |      |
|                       |      | /organism="Homo sapiens"   |                                      |            |                         |             |   |      |
|                       |      | /db_xref="Gene:380137"   |                                      |            |                         |             |   |      |
|                       |      | /db_xref="taxon:9606"  |                                      |            |                         |             |   |      |
|                       |      | /clone="IMAGE:294626"  |                                      |            |                         |             |   |      |
|                       |      | /feature_1fb="Soares fetal liver spleen INTLS"   |                                      |            |                         |             |   |      |
|                       |      | /sex="male"  |                                      |            |                         |             |   |      |
|                       |      | /cov_stage="20 week post conception fetus"   |                                      |            |                         |             |   |      |
|                       |      | /lab_host="DH10B (ampicillin resistant)"   |                                      |            |                         |             |   |      |
|                       |      | /note="Jordan Liver and Splein Vector: pTZ19 (Pharmacia)   |                                      |            |                         |             |   |      |
|                       |      | with a modified polylinker site 1: Pac 1 Site 2: Eco RI:   |                                      |            |                         |             |   |      |
|                       |      | 1st strand cDNA was ligated with a Pac 1 - oligo(dT) primer  |                                      |            |                         |             |   |      |
|                       |      | [5' AACCTAGCATTTATTAAGAATTGTTTTTTTTTTT 3']   |                                      |            |                         |             |   |      |
|                       |      | (Pharmacia); digested with Pac 1 and cloned into the Pac 1   |                                      |            |                         |             |   |      |
|                       |      | and Eco RI sites of the modified pTZ19 vector. Library   |                                      |            |                         |             |   |      |
|                       |      | went through one round of normalization. Library   |                                      |            |                         |             |   |      |
|                       |      | constructed by Bento Soares and M.Fadima Honrado."   |                                      |            |                         |             |   |      |
| BASE COUNT            |      | 116 a 120 c 100 g 118 t  |                                      |            |                         |             |   |      |
| ORIGIN                |      |  |                                      |            |                         |             |   |      |
| Query Match           |      | 10.0%  | Score: 214.85                        | Dh: 11     | Length: 454.            |             |   |      |
| Best Local Similarity |      | 99.1%  | Prod. No.: 1.4e-41                   |            |                         |             |   |      |
| Matches 216:          |      | Conservative   | 0                                    | Mismatches | 2                       | Indels      | 0 |      |
| Gy                    | 1929 | aagatcccaatcaaatgctcagaatgctggagagtaacaaatgacaatataacatgaa   | 1988                                 |            |                         |             |   |      |
| Dd                    | 7    | aaatccacaaiaaagtgcacagctgcgaacgacacaaacaaacaaacaaacaaac  | 66                                   |            |                         |             |   |      |
| Gy                    | 1989 | gattcataaacacataaacacatatattatctctatgactcttctatgactctctaa  | 2048                                 |            |                         |             |   |      |
| Dd                    | 67   | gattctatcccacaaiaacacacattctctctctctctctctctctctctctctct   | 126                                  |            |                         |             |   |      |









